DOCUMENT RESUME

ED 076 220 LI 004 333

TITLE	National Serials Data Program; Phase I, Final Report.
INSTITUTION	Library of Congress, Washington, D.C. Information Systems Office.
PUB DATE	Jun 69
NOTE	94p.; (3 References)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Automation; *Data Bases; *Library Automation; Library Surveys; National Libraries; Periodicals; *Serials; Systems Development; Use Studies

ABSTRACT

The primary objective of the National Serials Data Program is the design and implementation of a national serials information system. The goals of this system are the provisions of timely, accurate, and comprehensive information about serial publications within a framework of quantitative efficiency and agreed upon cost effectiveness. The intent of the program is the creation of a mechanized data base which would encompass all the world's serial literature without regard to subject or disciplinary classification. The data base would contain information relating to the description and location of all serials in order to increase the bibliographical and physical accessibility to serials. The three national libraries accepted the responsibility for implementing the first phase of the Serials Data Program. The Library of Congress was the executive agent for this project, and its Information Systems Office had the responsibility for project direction. The Joint Committee on the Union List of Serials acted in an advisory capacity. This document reports upon phase I of the project including results of a user survey and alternatives to the proposed system. (A related document is ED 063 009.) (Author/SJ)

FILMED FROM BEST AVAILABLE COPY

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 076220

NATIONAL SERIALS DATA PROGRAM

PHASE I FINAL REPORT

INFORMATION SYSTEMS OFFICE
LIBRARY OF CONGRESS
WASHINGTON, D.C.

June, 1969

LI 004 333

TABLE OF CONTENTS

I. INTRODUCTION

II. USER SURVEY

III. ALTERNATIVES

IV. CONCLUSIONS

Appendix I — DATA ELEMENT LIST (Introd. + Table of Contents only)
Appendix II — FORMAT (Not included here. Has been superseded by revised format)

I. Introduction

I. BACKGROUND OF THE SERIALS DATA PROGRAM

Serial literature is one of the most important parts of a library's collection and is indispensable to research in the fields of science and technology, social sciences, and humanities. Control of this type of publication has been of long standing concern to libraries and to scholarly and research communities. Serial literature is difficult to control because its data elements are highly mobile and require constant updating, and because titles are often difficult to identify, describe, and locate. Computer technology offers new opportunities and different approaches for controlling this changeable and voluminous literature.

Many libraries have experimented with automated techniques in handling their own serial literature, and this work has created an interest in the concept of national networks for the exchange of data. There has been an increasing demand to create a central store of data concerning serials that would be available to all.

The federal government's interest in creating a serials data program began with the concern of the National Science Foundation and the Committee on Scientific and Technical Information to bring about more control over scientific and technical journal literature. COSATI'S interest in scientific journal problems grew out of a proposal made by Dr. Stafford L. Warren, Presidential Special Assistant for Mental Retardation, to create a computer based central pool of all scientific literature. A Task Force for the Study of Scientific and Technical Journals was established within COSATI in April 1964, and was chaired by Dr. Burton W. Adkinson.

In May 1964, Dr. Adkinson suggested the subject of a machine manipulable inventory of journals in science and technology as a topic for further study by the COSATI Task Force. That same month a special Working Group on the National Inventory of Journals in Science and Technology was established as a Subcommittee of this COSATI Task Force to consider the need of such an inventory.

During the summer of 1964, this Working Group addressed itself to such problems as: a definition of a serial; an estimate of the population of the scientific and technical serial universe; the procedures in starting an inventory; what functions an inventory might perform; and what benefits could be derived from such an inventory. There was much discussion and effort directed toward establishing the feasibility of using the serial records of the three national libraries as a starting point upon which to base a journal inventory. An alternate proposal, advanced by the Library of Congress, was to mechanize the World List of Scientific Periodicals. These two approaches were the subject of much debate during August 1964.

In September, the Working Group suggested to the Task Force that an outside study effort would be required. When the Task Force endorsed this recommendation, NSF was asked to support the study, and in October, it solicited proposals from private industry for a study of the feasibility of creating a national inventory of the world's scientific and technical serial publications. Proposals were sent to NSF during November 1964, and in April 1965, the contract was awarded to Information Dynamics Corporation.

The IDC feasibility study was submitted in December 1965*. Its basic conclusion was that the development and operation of a scientific and technical serials data program was technologically and procedurally feasible. The study also enumerated the benefits to be derived from such a program and presented cost estimates.

In the IDC report, it was suggested that further progress toward development of a serials data program would require the assumption of responsibility by some organization or organizations for its operation and funding, and that systems requirements, complete systems design, and a detailed operating plan should be prepared before operations began.

The IDC report was submitted for review to a number of organizations and individuals. Many reviewers agreed in principle with the overall recommendations but disagreed about specific aspects of the proposed program.

In January 1966, Dr. Adkinson asked the Librarian of Congress for an expression of the Library's view of the proposed National Inventory of Scientific and Technical Serials as outlined in the IDC report. In replying to Dr. Adkinson on January 21, 1966, the Librarian stated that the Library agreed with the principal findings of the report but suggested the need of a market survey and the need to extend the coverage to all serials. The Library expressed its interest in assuming the responsibility of a serials data program if the proper funding were provided; however, it was made clear that the Library could not make a commitment at that time.

---

*Information Dynamics Corporation. A serials data program for science and technology. Prepared by William A. Creager and David E. Sparks. Reading, Mass., 1965. 190 p.

Dr. Adkinson then queried the Association of Research Libraries as to its interest in creating an inventory of scientific and technical serials. At the 1966 Conference of the American Library Association, ARL appointed an ad hoc committee to work on the serial problem, with the understanding that an inventory should be broadened to include serials in all subject fields and that effort should be made to produce it in machine-readable form. This ad hoc committee subsequently became the Subcommittee on World List of Serials of the Joint Committee on the Union List of Serials.

The Subcommittee in turn invited the Library of Congress to consider the problem. The Library submitted its recommendations to the Subcommittee at a meeting held at the Library on December 9, 1966. At this meeting, the Subcommittee asked the Library to prepare a proposal in time to be submitted for consideration by the Joint Committee on the Union List of Serials at the American Library Association Midwinter Meeting in January 1967.

At the 1967 American Library Association Conference, the Directors of the National Agricultural Library and the National Library of Medicine and the Librarian of Congress announced that the three national libraries would undertake the Serials Data Program on a cooperative basis. The proposal was accepted in August 1967, and work began on the program.

The total program has been projected in four phases:

Phase I - Preliminary Design
Phase II - Reduction to Practice
Phase III - Pilot Project and Planning for Large-Scale Conversion
Phase IV - Conversion and Implementation of the Total Program.

This phasing is essential because practical experience gained in each phase is of great value in determining directions that subsequent phases should take and ensuring that concepts are sound and procedures well-tested before large expenditures of money and manpower are committed.

The three national libraries accepted the responsibility for implementing the first phase of the Serials Data Program. The Library of Congress was the executive agent for this project, and its Information Systems Office had the responsibility for project direction. The Joint Committee on the Union List of Serials acted in an advisory capacity.

The major objective of the program is the creation of a mechanized data base which would encompass all the world's serial literature without regard to subject or disciplinary classification. The data base would contain information relating to the description and location of all serials in order to increase the bibliographical and physical accessibility to serials.

PHASE I - OBJECTIVES

The first phase had as objectives the definition of serials, the data elements needed to control them, and the development of a content format for serials. Because all serials were to be included in the projected national system, a firm definition would have to be determined in order to determine which classes of material had to be included and, as a corollary, how many had to be included. Based upon the definition and objectives of a national system, those data elements could be selected which would carry out the objectives. Because objectives were to be accomplished at least cost, those items and elements were to be included which would be most useful to users.

I. TASK A. - SERIALS DATA ELEMENT PROJECT

Purpose

The purpose of this task was to develop a comprehensive set of elements necessary for the identification, description, and location of serials. This set of data elements would be the basis for developing a content machine format for serials. The approach was to identify all possible data elements associated with the bibliographic identification, location, and control of serial literature and to recommend possible inclusion or exclusion of each data element on the basis of: a) usefulness in a serials system; b) usefulness in machine processing of data; and c) the cost of adding the data element to the record. All categories of serial literature were to be reviewed in this analysis, and previous work in this field was to be utilized and supplemented where needed.

Progress and Results

A. Review of Existing Work

1. Review of Other Automated Serials Systems

Work on Task A began with reviewing developments and accomplishments of other existing serials automation projects. At the outset of the task, all institutions involved in some stage of serials automation were identified. Approximately 300 institutions were contacted and asked to provide a description of their system, input formats, samples of output, any discussions of problems encountered, cost figures if available, and description of future plans. The response to this inquiry was high. More than 250 institutions sent materials, and unsolicited material was received during the ensuing year. From this material, some 200 formats were analyzed in depth. Data elements were extracted from each of these in a matric-type array showing: 1) data elements used; 2) name of the institution using them; 3) type of serials system (full automated checkin, acquisitions, union lists, etc.); 4) codes used with the data elements; 5) any problem associated

with their use; 6) how the data elements were used -- whether in fixed or variable field; and 7) any costs involved in providing the data element.

The results of this review were somewhat disappointing. It was found that previous work in the field did not yield as much information as had been expected. Although the automated systems were working satisfactorily, they were limited in size and scope. To handle a national system such as the Serials Data Program, much more work would need to be done, and much of this endeavor would be pioneering work in the strictest sense. Unfortunately, many areas and problems which are of concern to a national program were not touched upon in any of the systems reviewed.

The most frequent use of automation or mechanization in the serials systems survey was the production of a title list with simple, summarized holdings. Sometimes additional information such as call number, library shelving location, and subject was also given. As an outgrowth of these individual library lists, area union lists of serials were being produced. At the time of the survey, 72 institutions were producing title lists. No attempt was made to completely survey those working toward developing statewide or regional union lists as this activity was just beginning to be widespread at the time the survey was made (August 1967).

The field of acquisitions was the next most popular use of automated or mechanized equipment. Fifty-nine institutions sent formats and descriptions of their work in automating the accounting functions of serials. Many institutions started with this kind of activity and then planned to expand into other areas of serials control. Initially, these acquisitions systems were not tied to any automated check-in procedures but were used to provide control over the renewal and accounting functions of serials subscriptions. Journal circulation systems were found to be operational in 26 institutions. These institutions were predominantly special libraries and information centers were rapid routing of serial material was a necessity. A few libraries (7) had mechanized or automated only their binding functions for serials.

Approximately 30 institutions were found to have some method of automated check-in control for serials. Most of these check-in systems were variations of the "computer produced expected arrival card" type of check-in system first developed by the University of California, San Diego, in 1961. They varied in sophistication, some including check-in acquisitions, binding, claiming, and routing into one module. Not all of these systems were operational. Some had developed problems when trying to handle more than 15,000 serials with this type of check-in control.

The survey found only seven institutions that were developing more sophisticated on-line systems. Only one of these was operational; the others were still in the specification and developmental stage. Very few systems were attempting to provide full descriptive information about serials. Only one system was giving full descriptive information on a small scale basis and that system is no longer being maintained. None of the automated systems surveyed attempted to provide full cataloging information as well as control information for serials.

2. Review of Serials Cataloging.

Traditional serials cataloging rules and practices were reviewed with the purpose of extracting the data elements used. Four forms of cataloging were analyzed: 1) the ALA form with entry under latest title and cross references from preceding ones; 2) the Anglo-American code, with successive entries with changes in title; 3) the title/author rather than corporate entry approach; and 4) the title as it appears on the piece. This study added more data elements to the data base gathered in the review of other automated serial systems.

3. Review of Special Studies.

Studies pertaining to specific areas related to serials were also analyzed. Some of these included special studies on the frequency patterning of serials, on abbreviated titles used for serials, and on any numbering or coding schemes developed for serials. The report of the Special Project on Data Dlements for the Subcommittee on Machine Input Records of the USASI Z-39 Committee* was used as well as the IDC report.

4. Summary of Review.

This review provided a base of data elements upon which to start work on compiling a comprehensive list. However, a great deal of refinement and augmentation of these data elements was needed. Refinement consisted of condensing the data elements into acceptable terminology. A semantic problem existed in that many times the same data element was called by different names. Since the data base reflected only what other systems had used and those elements used in traditional cataloging, it was necessary to take a "utopian" approach and set down all possible data elements needed for the description and control of serials. This, of course, added more data elements to the data base.

*Curran, Ann T., and Henriette D. Avram. The Identification of data elements in bibliographic records. 1967.

B. Data Element List

At this point, a data element list was compiled. For each data element the following information was given: 1) a definition as to what was meant by the term; 2) examples and codes to clarify the definition and to give an indication of how the data element could be handled; 3) problems in the use of the data element or alternative means by which it could be handled; and 4) whether the data element should be considered for national use or for strictly local, internal use. (Throughout the compilation of data elements, the emphasis was on national elements, i.e., those that pertained to a serial regardless of where it was held.) The first data element list contained 278 data elements. It was a comprehensive, almost "utopian" list since the approach was to define all that was necessary and useful for the control and description of serials and then to cut and refine. It was this list that was used in the user survey conducted by Nelson Associates. The examples and codes listed with the data elements are not all-inclusive but are meant to clarify the definition. The "local" data elements are strictly indicative of data elements used by other systems. They do not reflect the data elements needed by the Serials Data Program or the Library of Congress for local control, nor are they tailored for any particular library's local needs. They are given as an indication of what a local system can use and to elicit information on how much space should be left for "local option". The list of data elements is presented in Appendix I.

C. Other Subtasks in Task A.

1. Categorization of Serials.

One of the subtasks in this first task was to define categories of serials explicitly and estimate the proportion of serials that fell into each category. Fifty-six categories of serials were defined (e.g., scholarly periodicals, house organs, technical reports, press releases, etc.). A 6,000 record sample from the Library's serial record was used to test out this categorization. At the conclusion of the test, it was determined that to categorize all serials precisely required many subjective decisions as a clear-cut division between types of serials is not always possible.

2. World-Wide Population of Serials.

Using union lists, national bibliographies, UNESCO figures, and other available data, an attempt was made to estimate the world-wide total of serial titles, both current and retrospective. Again, the problem of definition of a serial hampered this study since no one source could be used and each source defined serials in a different manner. An estimate was made of 1,300,000 serials.

3. Universal Numbering Scheme for Serials.

Early in the program, it was recognized that a numbering or coding scheme to identify each serial title uniquely would be necessary.

Some preliminary work on developing such a scheme was undertaken, but because the ramifications of such a scheme would be so widespread, it was decided to ask the USASI Z-39 Committee to set up a subcommittee to study this problem and develop a scheme. In August 1968, Subcommittee 20 (Identification Coding of Serials) was formally constituted.

4. Serial Record Sample.

As part of Task A, a sample of approximately 6,000 records was drawn from the Library's serial record to ascertain file size and file characteristics. The analysis of the sample was completed in August 1968, and in addition to file size, it gave statistical data such as the number and proportion of types of records, languages, currency vs. non-currency of holdings, frequency, etc. File size of the serial record was estimated at 640,992 entries. This figure includes all types of entries such as cross references, authority cards, etc., as well as check-in records. Of this number, 185,381 are estimated to be current serial check-in records.

II. TASK B. - USER SURVEY

Purpose

The purpose of the second task was to assess the interest of potential clients in the Serials Data Program, to elicit comments, suggestions, and criticism of the data elements as set forth in Task A, and to determine the nature of consumer demand with respect to the desired outputs, products, and services that could be provided from the program.

Progress and Results

Using the data element list compiled as part of Task A, Nelson Associates designed a questionnaire which was presented to 168 respondents. These respondents represented the broad spectrum of personnel handling serials in a variety of libraries or library-related organizations. Each person receiving a questionnaire was also interviewed by a member of the survey team from Nelson Associates so that respondents had an opportunity to elaborate upon their responses as well as to discuss problems, criticisms, or suggestions relating to any of the data elements.

The final report from Nelson Associates consisted of a brief description of the survey, a summarization of the questionnaire tabulations and all comments made by the respondents relating to any of the questions or date elements. In addition, a complete computer tabulation showing responses to the questionnaire both by total respondents and by type of respondents was made available to the Library.

The survey provided ample data and many useful suggestions and comments about the data elements or the program in general. Unfortunately, the responses to some of the important questions were inconclusive. For example, there was no evidence that a standard serial definition was applied uniformly by all the respondents or that a definition could be formulated which would satisfy a majority of users. From their comments, it appeared that most respondents were more willing to accept a broad rather than a narrow definition.

With regard to language representations in the Serials Data Program, 142 respondents (86.1%) indicated that publications in non-Roman alphabets, non-alphabetic languages, etc., should be entered in romanized form. The general feeling was that they would prefer to have the data in romanized form than not to have the data at all. Respondents also emphasized the necessity of having standard romanization schemes for this purpose.

A number of possible products and services of a national serials data program were listed in the questionnaire so that respondents could indicate their preferences. It should be noted that this discussion was based on the assumption that everyone would have some kind of access to a computer; therefore, the results should be evaluated with this in mind. The following table contains a list of some services or products which were rated as "high priority items" by the respondents.

Product or Service	No. of Respondents Ranking
Reports on new serials	150
Reports on changes in serials	148
Reports on ceased items	142
National union list of serials	134
Cataloging data on current serials	132
Cataloging data on ceased serials	107
Retrieval from file of any or all data on a serial or serials by selected data elements	98
Registration number for every serial	89
Directory of participating libraries	83
Abbreviation lists	75

The majority of products and services which would result from a national program are similar to those currently in existence. It was hoped that such outputs would be made available much more rapidly and would be more flexible than they are at the present time.

III. TASK C. - ANALYSIS AND PRE-DESIGN

Purpose

The purpose of Task C was to use the original data element list from Task A and the requirements expressed in the user survey from Task B to develop a realistic and definative format for serials in machine-readable form. Suggestions were received from both the three national libraries and the designers of the MARC II Format. The draft serials format was used as the basis of a test which provided statistical data for the cost model, such as frequency of use of a data element, character count per data element, number of characters per record, etc. This test of the draft format also provided improvements to the format and more insight into problems.

Progress and Results

A. Pre-Design

The serials processing at the Library was analyzed in depth. Using material furnished under contract by the United Aircraft Corporate Systems Center, functional flowcharts of the serials system were developed. The manner of logic development and expression was generalized with emphasis placed on commonalities with other systems, thus resulting in broad applicability of the logic. Files containing information about serials were analyzed, local data elements were collected, and information needed for the pre-design of an automated serials system to support the national data bank was collected.

B. Cost Study

The subtask of designing a cost-benefit model for the program was made difficult by the lack of firm information concerning the total utilization of the system. The user survey had not been developed with this purpose in mind so that the results could not be used in the construction of the model. The following discussion covers the factors considered in determining costs for the program.

A national serials system should provide the greatest possible precision to the largest number of users. Since such precision and quality of services increase the system cost, costing became a critical factor in this phase of system design. Benefits were to be provided with the least amount of expenditure, or, to express it differently, a system should be selected out of the various ones possible which had the highest benefit/cost ratio. A first step in this direction is to determine what costs should be considered within the scope of the system.

The scope of a system ideally includes the same area as that which is included in costing. Information handling systems, more than

other kinds of systems, characteristically influence and are prejudiced by related systems and sub-systems. Because boundaries cannot be precisely defined, conventions must be introduced in order to determine cost. It is emphasized that these conventions are directly related to costing and are only indirectly related to the intrinsic coverage of the system.

The broadest convention includes functions which would be necessary to feed bibliographic data into the system, facilities to store the data for immediate access at points of use, and communications necessary to update these storage facilities. It was recognized that a convention as liberal as this would produce a high total cost. This proved to be the case, and the resulting estimates were primarily useful in stimulating discussions of how comprehensive the National Serials Data Program should be. The result of these discussions are reflected in costs which are presented in Table I in IV. ALTERNATIVES.

A second convention includes in national costs the same functions required to feed and update a national bank together with facilities required to store the data at a limited number of regional centers. A national system might, therefore, include regional centers in addition to a national center. The cost will not greatly change if the number of regions were estimated to be around nine. Present technology would support the number of accesses provided unchanging information did not have to be distributed to all users by the national system.

The third convention includes only those functions and facilities to store information at the national center. A single memory has been postulated which could accept updating information from a limited number of processing locations and from a large number of individual terminals. It would be limited in the number of accessing points it could store. Although a "national" center seems to imply a single physical memory, it is only slightly more difficult to divide the store among a limited number of interacting memories.

The scope of the national system is further characterized by the functions which should be included. The National Serials Data Program will hold its entire store in machine-readable form although a number of stores are available at different costs per record. A convention can be adopted by which the national system costs include only those which are attributed to the central responsibility. Printed cards and machine-readable bibliographical materials present an analogy. The national system in these cases provides data but does not supply means of access. Costing of an analogous serials system would, therefore, include only the functions of putting data into the store and maintaining it. Retrieval would not be charged to the national system, although a

retrieval facility would of course be implemented. By this convention, users would provide their own retrieval facilities or would reimburse the national system for retrieval services.

A study was made of the details of cost experience in similar systems. This study indicated that even with a limited number of sites, retrieval costs would be approximately equal to the total of all other costs. This conclusion results from the fact that many of the elements for retrieval are proportional to the number of file entries and to the number of queries. It can be assumed that a file as large as this will be addressed by means of inverted files. An inverted file concept involves greater input and storage costs. It offsets these against time and cost savings at the time of retrieval. Relationships among the various functions to be considered in retrieval show that retrieval without inverted files is linear; however, when inverted files are used, time, number of retrievals, and complexity of retrievals interact, the exact interaction depending upon the design of the system. Hence, the cost of retrieval is directly proportional to the number of inverted file entries created, to the size of the files to be searched, to the number of retrieval calls on the file, and to the number of changes to the master file which indirectly affect inverted files. Because most of these are the elements which cause other processing costs to vary, it follows that retrieval will vary with the other functions.

Costing to the individual data element in an automated system is complicated by several facts. Historically, costs of determining and recording bibliographic data have not been collected to this level. They must, therefore, be derived from larger functions. Secondly, there is not sufficient experience to prove comparability between costs which apply to a manual and to an automated version of the same operation. Thirdly, certain of the factors required by a cost model can be obtained only from the experience of people who have been engaged in serial work, and it could not be proved that this experience covered the entire population of serials. Estimates are sufficiently valid, however, so that costs can be computed for areas of elements rather than for individual elements. Because historic costs and experience apply to areas of work rather than to individual data elements and because man and machine functions can be more closely associated than can individual elements, these costs are also more valid.

PART I. SUMMARY OF THE USER SURVEY

The final report of the user survey conducted under contract was detailed in its analysis. It consisted of a brief summary of the survey, a summarization of the questionnaire tabulations, and comments made by the respondents relating to any of the questions or data elements. In addition, a complete computer tabulation showing the responses to the questionnaire both by total respondents and by type of respondent was made available to the Library. Conclusions were not required under the contract nor were they drawn since it was felt that the results constituted another input to decisions which would have to be made in regard to the entire program.

Tabulations of findings relating to national data elements and possible outputs of the serials data program are presented here and comments by the respondents are summarized. No essential information contained in the contractor's report is omitted. For clarity of presentation, some questions are paraphrased and for others the topic is indicated rather than the question stated.

PART II. FINDINGS

A few of the general points brought out in the results of the survey are worthy of mention here. One such finding was an expressed concern that a national program be designed to take maximum advantage of the characteristics of a computer system, rather than be based on present manual standards. One of the most critical problems relating to this was considered to be the updating of the national data bank. The need for obtaining a currently updated product was felt to be so important that respondents indicated that they would sacrifice other things to obtain it.

An overriding concern about government publications was evident throughout the report. Also, many people emphasized the need for comparative cost figures.

As was stated in the discussion of the background for the survey, people were encouraged to comment freely, and they did so. This resulted in much useful information. These comments, criticisms, and suggestions have provided useful data in constructing the Serials Data Program format.

PART III. TERMINOLOGY

Defining Terminology: One part of the questionnaire was formulated as an attempt to determine the definition of serials being used by the various respondents and to identify the specific types of publications which should be considered to be serials for the purposes of the National Program. As would be expected, agreement upon one standard definition did not result.

Definition of a Serial: Of those respondents who stated that they used a codified definition of a serial, 55% used the ALA definition: [a serial is] . . . a publication issued in successive parts, usually at regular intervals, and as a rule, intended to be continued indefinitely." Of the remaining group, there was a tendency toward a broader definition.

Type of Publications Considered Serials and Whether to be Included in NSDP: Fifty-two types of publications were presented. In general, it was felt that whether a given publication should be included in the Serials Data Program was a pragmatic consideration and was related to whether the publication would be entered in the MARC monograph format or in the Serials Data Program. The following table, in order by rank, shows the percentage scores for each publication type.

Rank	Type of Publication	Percent of Total Respondents Who Indicated "Serial should be included"
1 (Tie)	Scholarly journals (periodicals)	99.4%
1 (Tie)	Professional publications	99.4
3	News publications (trade magazines)	96.3
4	Abstracts	95.7
5	Popular publications (mass media)	95.7
6 (Tie)	Indexing services	95.0

Rank	Type of Publication	Percent of Total Respondents Who Indicated "Serial should be included"
6 (Tie)	Organs	95.0%
8	Yearbooks	94.4
9	Book reviews	93.1
10	Conference proceedings	92.5
11	Official gazettes	91.4
12	Little magazines	90.1
13	Statistics	83.5
14	Indexes	80.9
15	Law reports and digests	80.1
16	Annual reports	77.5
17	Almanacs	76.3
18	Alumni publications	75.9
19	Bibliographies	74.8
20	Government organization manuals	73.0
21	Technical reports	71.2
22	Monographic series	70.8
23	Biographical dictionaries	69.2
24	Newspapers	66.7
25	Directories	64.2
26	Registers	58.2
27	House organs - external	55.3

Rank	Type of Publication	Percent of Total Respondents Who Indicated "Serial should be included"
28	Undergraduate publications	53.8%
29	Digests	51.6
30 (Tie)	City directories	50.3
30 (Tie)	Local history or genealogy	50.3
32	Catalogs	49.4
33	Film catalogs	46.5
34	Checklists	44.5
35	Press summaries	43.6
36	Handbooks	41.1
37	Calendars	39.1
38	Guidebooks	38.6
39	College catalogs, bulletins, and announcements	37.9
40	Pseudo-serials	36.1
41	Looseleaf publications	35.9
42	Program guides	33.1
43	House organs - internal	32.5
44	Exhibition catalogs	29.1
45	Collections of reprints	27.1
46	Trade catalogs	22.3
47	Telephone directories	21.0
48	Publishers' series	17.0
49	News releases	15.1

Rank	Type of Publication	Percent of Total Respondents Who Indicated "Serial should be included"
50	Programs for concerts, plays, etc.	15.0%
51	Dictionaries	11.8
52	Books-in-part	9.4

Categorization of Serials

Q: Should the system have the capability to retrieve publications by categories?

Essential	Useful But Not Mandatory	No Opinion	No	Number of Respondents
43.3%	40.2%	3.7%	12.8%	164

A number of people indicated that a limit should be placed on the number of cateogires used for this purpose.

Government Publications

Q: Should the program discriminate between government and nongovernment publications?

Yes	No	Number of Respondents
51.1%	42.9%	163

PART IV. GENERAL PROBLEMS RELATED TO THE SERIALS FORMAT

Cataloging Entry Forms

Q: Which of the following cataloging entry forms must the Serials Data Program provide for retrieval capability?

Form of Entry	Number of Respondents	Percent of Respondents
ALA Rules of Entry (entry under latest title)	110	66.3%
AA Rules (successive entries for changes in title)	110	60.2%
Title/Author rather than Author/ Title	32	19.3%
Title as it appears on the piece	100	60.2%

There was a great deal of comment on this question, indicating that the problem of entry is of prime concern. All comments reflected the opinion that access from all points is mandatory. Many respondents pointed out that computer processing negates the concept of "main entry". Data processing respondents commented that it is more feasible to have each title in a separate record [AA Rules approach] as it is easier to reformat from split entries to one record than the reverse.

Vernacular versus Romanization

Q: What method would you suggest for handling serials in non-Roman alphabets?

Do not include serials in non-Roman alphabets	4.1%
Enter such material in romanized (transliterated) form	86.1%
Enter in the vernacular	29.1%

In general, the comments showed that the respondents would like to have the material in the vernacular, but with the present state of the technology, wondered if it were worth the trouble. A major concern of the respondents was that a standard transliteration for each non-Roman alphabet must be agreed upon.

PART V. DATA ELEMENTS - TITLES

The following table summarizes the responses for one group of titles.

Element	Full Title	Short Title	Sub Title	Main Series Title	Subseries Title	Analytic Title	Distinctive Title	Title As it Appear on the piece
Number of Respondents	155	147	148	159	159	158	158	159
MANDATORY - generate associated elements for cross references and translations	74.8%	51.0%	18.2%	73.6%	53.5%	41.1%	46.8%	56.6%
MANDATORY - and generate associated elements for cross references only	5.8	15.0	9.5	12.6	14.5	11.4	11.4	11.3
MANDATORY - generate associated elements for translations only	.6		.7				.6	
MANDATORY - do not generate associated elements	2.6	8.2	16.9	7.5	11.3	11.4	11.4	6.9
Desirable	7.7	10.2	29.1	1.9	11.3	19.6	17.7	12.6
No Opinion	3.9	5.4	8.1	4.4	6.9	5.7	6.3	3.1
Not Desirable	4.5	10.2	17.6		2.5	10.8	5.7	9.4

A number of respondents expressed difficulty in understanding what was meant by "generate all associated elements for cross-references and translations;" however, they did agree upon the importance of the above titles. The respondents generally felt that portions of the titles should be delimitted and that access by short title and subtitle was wanted. In regard to the other titles listed above, respondents indicated the desire to query the system under any one of these and thus to be able to obtain the full available information.

Another grouping of types of titles and the responses follows:

	Mandatory	Desirable	No Opinion	Not Desirable
Abbreviated Title	35.6%	36.3%	9.4%	18.8%
Uniform Title	30.9	20.4	30.9	17.8
Added Title Page Title	47.2	31.7	13.7	7.5
"At Head of Title" Title	21.5	38.0	15.8	24.7

Some specific questions relating to these titles warrant comment.

Abbreviated Titles

Q: Do you feel that the USASI Title Word Abbreviation for periodicals is adequate to cover the entire range of serial publications and still remain an intelligible and unique set of codes?

Yes	No	No Opinion	No. of Respondents
28.6%	51.6%	19.8%	161

Respondents differed on whether a code must be unique. Some felt that the abbreviated title was a "waste of time," and that this whole area needed more attention. The comments indicate that many are confusing "abbreviated title" and a "standard serial code" and would prefer the latter. In general no agreement was expressed.

Uniform Titles

Q: Must the entry of uniform titles be allowed to fit local practices?

Yes	No	No Opinion	No. of Respondents
27.3%	44.7%	28.0%	161

Many respondents questioned the existence of such titles for serials. A number of comments indicated that this would be better as a local element.

Additional Titles

The response favoring the inclusion of a third group of titles is tabulated below.

Q: Should the program have the capability of retrieving these titles (see below)?

Yes	No	No. of Respondents
69.6%	30.4%	158

Q: Which ones should be included? (This is based upon 109 respondents having answered the question.)

Type of Title	Percent of Respondents
Cover Titles	95.4%
Masthead Titles	57.8
Running Titles	45.9
Spine Titles	56.9
Caption Titles	30.3
Added Titles	67.0

The comments made showed that all of the above listed titles should be included provided variations were substantial. It was felt that the judgement of the cataloger should be the basis for inclusion.

Translated Title

Q: If a translated title exists at all, . . . should this be a separate data element?

No. of Respondents	164
Mandatory	82.3%
Desirable	14.0
No Opinion	2.4
Not Desirable	1.2

General Comment: This was a poorly worded question and caused some confusion among the respondents. This factor should be borne in mind when interpreting the results.

PART VI. DATA ELEMENTS - CONTROL NUMBERS

Universal Numbering Scheme

Q: Is a universal numbering scheme for serials needed for this program?

Mandatory	Desirable	No Opinion	Not Desirable	No. of Respondents
48.2%	39.8%	8.4%	3.6%	166

Many respondents felt this to be one of the most important items in the program. They saw the use of a universal system as a means to reduce the problem of variations in entry and as a way of linking records to one another. On the other hand, others brought out that there would be some difficulty in developing such a scheme.

Comments relating to the specific type of system included suggestions for an arbitrary registration type number, a structured code with bibliographical meaning, and a code derived by the machine from the entry. A number of people encouraged the development of a completely new system, as soon as possible.

Use of CODEN as a Universal Numbering Scheme

Q: Is CODEN flexible enough to be used as a base for the universal numbering scheme?

Ninety-six out of 163 respondents (58.9%) said that they were familiar with CODEN. Eighty-seven of these people answered the above question.

Yes	No	No. of Respondents
47.1%	53.9%	87

Many of those answering "yes" added that they did not really favor the use of CODEN in this context. Others that did answer "yes" qualified their answering by saying they really had no opinion.

Other Control Numbers

Q: Short of a Universal Numbering Scheme, which of the following merit use as a data element in the Serials Data Program?

	Ought to be Included	Do Not Include	No Opinion	No. of Respondents
CODEN	51.6%	12.4%	35.9%	153
LC Card Number	75.0	12.8	12.2	156
National Bibliography Numbers	45.8	22.9	31.4	153
PL-480 Numbers*	22.9	36.6	40.5	153
LACAP Numbers*	12.6	42.4	45.0	151
Subscription Agent Number	7.3	59.3	33.3	150
Local System Numbers	25.2	50.3	24.5	155

* This low response is probably due to the fact that few of the respondents were participating in these programs.

Respondents pointed out the fact that none of these systems control all titles. Some felt that only the LC card number should be kept and that all the other systems should be eliminated. However, others suggested that any control number with wide applicability should be included in the Serials Data Program.

PART VII. DATA ELEMENTS - ISSUING BODIES AND AUTHORS

Personal and Corporate Authors

Q: Should all personal and corporate authors be retrievable?

	Percent Positive Response
All should be retrievable	12.8%
Should be a limit to these	87.2%

Number of respondents: 164

A number of respondents emphasized the need for full format for each author identified. A suggestion was also made to use a thesaurus structure for some of this information, e.g., a publisher's thesaurus, so as to prevent cluttering the system.

Personal Authors

Q: Should the following subfields be included as data elements for personal authors?

Data Element	Percent Positive Response
Relator*	90.1%
Identifier	69.7
Title Designating Rank	16.2

*Complier, editor, translator

Number of respondents: 142

Some respondents felt a delimiter was needed, but that the subfield need not be retrievable. It was felt that these subfields should be included if they were necessary to identify a personal author, and that the Anglo-American Cataloging Rules, MARC procedures, and the like, should be followed.

Corporate Authors

Q: Should there be the capability of noting whether the corporate author and issuing body is the same?

Yes	No	Number of Respondents
73.4%	27.6%	163

The suggestion was made that two slots be available and that both be filled in only when the information is different. Then, if the issuing body were missing, it would be by implication, the same as the corporate author. Some respondents indicated that query should be possible without specifying category.

Issuing Bodies

Q: Should issuing bodies be classified as to type?

Mandatory	Desirable	Not Desirable	No Opinion	Number of Respondents
14.6%	53.7%	23.8%	7.9%	164

Some respondents commented on the usefulness of this information. Others pointed out the difficulties which would arise in attempting to classify issuing bodies.

Q: Which of the following data elements describing authors and issuing bodies should be entered separately in the Serials Data Program?

Data Element	Percentage of Total Respondents			
	Mandatory	Desirable	Not Desirable	No Opinion
*Name of Applicable Conference or Meeting	84.5%	9.9%	4.3%	1.2%
*Body of Which Serial is the Official Organ	74.8	18.9	4.4	1.9
*Joint Issuing Body	72.3	20.8	3.8	3.1
*Former Issuing Body	69.2	19.5	6.9	4.4
≠Date of Conference Held	64.8	17.3	6.2	11.7
*Sponsoring Agency of Conference	60.8	31.0	5.7	2.5
*Associated Corporate Authors	58.8	25.6	10.0	5.6
≠Place Where Conference Held	43.4	36.5	8.2	11.9
≠Date of First Conference Held	48.4	25.8	13.2	12.6
*Associated Personal Authors	43.0	28.5	19.0	9.5
*Other Bodies Associated With Serial (e.g., Preparer, Compiler)	35.7	40.8	11.5	12.1
*Institution Where Conference Held	32.9	47.5	7.6	12.0
≠Reference to Previous Meetings of the Same Group	20.8	41.5	13.8	23.9

Total Respondents - Approximately 160

* Respondents indicated these data elements should be retrievable.

≠ Respondents indicated these data elements should be available for information only.

Some respondents said that these data elements would be useful for reference operations, others pointed out that they change, even when the title does not.

Most respondents felt that these elements should be translated only if cataloging rules so require.

Translate	Do Not Translate	Number of Respondents
70.1%	29.9%	157

Q: How many joint issuing bodies should be entered?

Three or Less	More than Three	Respondents
83.2%	16.8%	125

Respondents noted that the question was poorly worded. Some would list all; others would use judgement for each case.

PART VIII. DATA ELEMENTS - IMPRINT

Place of Publication

Q: Should there be a distinction between place of publication and publisher's address?

Yes	No	Number of Respondents
53.4%	46.6%	161

A number of respondents commented that the publishers address should be in a separate file and that the machine system should not be burdened with it.

Q: What subfields for publisher's address should be in Serials Data Program?

Subfield	Percent Mandatory
City	98.8%
Country	90.0
State or Province	84.4
Zip Code	75.0
Street Address	70.0

Total Respondents: 160

With regard to the above question, it should be noted that if there is going to be a separate publisher's file, it would be needed for acquisition and claiming purposes, so that all of this subfield information would be available.

Q: Should a publisher's code be developed?

No	Yes	Number of Respondents
57.3%	42.7%	157

Some of the respondents said that such a code should not be developed, but that it would be nice to have. Those who answered positively saw problems involved, and were not optimistic about its development.

Q: Should distributor or printer be identified, if different from publisher?

Yes	No	Number of Respondents
74.8%	25.2%	163

Some of those who were in favor of such an identification qualified the response by saying that they did not think it was mandatory. Several suggested distinguishing between printer and distributor.

Q: Should the editor's name be included?

Yes	No	Number of Respondents
56.3%	43.8%	160

Many who circled "no" added that it was, however, desirable (not a choice provided by the questionnaire). A number of respondents said that the editor's name should be included on a selective basis only.

PART IX. DATA ELEMENTS RELATING TO CONTENT

Classification Numbers

Q: Which of the following classification numbers might be included as useful data elements?

Data Element	Mandatory	Desirable	No Opinion	Not Desirable
Library of Congress Call Number	72.7%	17.9%	3.7%	3.7%
National Library of Medicine Call Number	37.7	24.7	28.4	9.3
National Agricultural Library Call Number	26.5	24.7	35.8	13.0
NAL Subject Category Number	17.3	14.2	51.2	17.3
Dewey Decimal Number (complete)	37.0	22.2	22.2	18.5
Dewey Decimal Number (short form)	6.3	18.4	33.5	41.8
Universal Decimal Classification Number	15.0	20.0	40.0	25.0
National Bibliography Classification Numbers	5.0	13.0	46.0	36.0
Superintendent of Documents Classification Number	35.4	18.0	26.7	19.9
Local Call Numbers	25.3	13.6	14.8	46.3

Number of Respondents: Approximately 160

Many respondents said that they would not use the call numbers of each of the three national libraries, but that these should be included. Commenting on the Dewey Decimal Classification number, respondents indicated that it should be included in principle, at least in the form of a three-digit number. Although the number of responden s expressing a need for the UDC number was low, some noted the need for this number abroad, and its merit for a computer system. Most of the comments relating to the local call number did not favor its use in the national record.

Subject Headings

Q: Which of the following subject headings should be included in the Serials Data Program?

	Mandatory	Desirable	Not Desirable	No Opinion
Library of Congress Subject Headings	70.7%	20.1%	3.0%	6.1%
National Library of Medicine Subject Headings	43.8	25.0	6.2	25.0
National Agricultural Library Agricultural/ Biological Terms	37.1	27.7	6.9	28.3

Number of respondents: varied from 159-166

Comments reflected some dissatisfaction with existing subject heading schemes and showed a desire for a more detailed and multi-subject approach to serials. A sample of other schemes suggested follows.

ASTIA [headings]
Chemical Abstracts [headings]
COSATI [headings]
Detailed multi-subject indexing of Serial
LEX [headings]
Ulrichs Periodical Director [headings]
NASA [headings]

Abstracting and Indexing Information

Q: Which of the following types of information should be indicated in the Serials Data Program?

Type of Information	Mandatory	Desirable	Not Desirable	No Opinion
Where indexed	59.8%	30.5%	3.0%	6.7%
Where abstracted	57.6	31.5	3.6	7.3
Where covered by citation indexing	49.1	33.9	3.6	13.3
Indication of whether abstracting and indexing coverage is partial or complete	50.9*		33.7	15.3

Number of respondents: varied from 162 to 165

* The questionnaire only provided the categories of "yes", "no", or "no opinion", for this question.

Language

Q: Should the program include elements specifying the language used?

Data Element	Specify Original Language	Specify Translated Language	Specify Both	Specify Neither	Number of Respondents
Language of Serial itself	21.0%	3.0%	75.4%	1.6%	167
Language of Table of Contents	8.1	9.4	40.0	42.5	160
Language of accompanying abstracts	11.0	19.0	54.0	16.0	163
Language of Title only	13.8	3.9	45.4	36.8	152

In general, the comments indicated a need for specifying the languages of each of the above on the theory that the library user wants to know whether he can read the serial, and if not, whether he can read any portion of it, e.g., abstracts.

Translations

Q: Should a code be included to denote the extent of translation (e.g., cover-to-cover)?

Yes	No	Number of Respondents
69.5%	30.5%	164

Q: Should a code be included to denote the original language for translated material?

Yes	No	Number of Respondents
94.6%	5.4%	166

Elements Dealing With Content

Q: Should a code be developed to indicate whether or not a serial has a stated editorial policy?

Yes	No	Number of Respondents
11.6%	88.4%	164

Most respondents indicated that this would be one of the least useful of possible elements.

Q: Should the program include a data element which would identify the nature of contents of a serial (e.g., whether it contains book reviews, advertising . . .)?

Yes	No	Number of Respondents
53.0%	47.0%	164

Some of those whose response was positive qualified their answer by saying that this information would be desirable but not mandatory (a choice which was not provided in the questionnaire). A number of people suggested that identification of the nature of the contents of a serial should be done selectively.

Q: Is it possible to classify all serials by their intellectual level?

Yes and		No	Number of Respondents
Practical	Not Practical		
12.0%	33.1%	54.8%	166

Comments reflected the feeling that the Serials Data Program should avoid subjective qualifications.

Limited Distribution

Q: Should a code be developed to identify material of a confidential or limited distribution nature?

Yes	No	Number of Respondents
75.0%	25.0%	164

There was some confusion in understanding this question, in that some respondents assumed that it applied to subscription information. The intent was to single out items that were indeed restricted and when this was brought out in the interviews, many respondents expressed some doubt as to whether this type of material should be included in the national program.

Edition Information

Q: To what extent should the Serials Data Program include the capability to retrieve foreign language editions?

Choice	Percent
Separate entries, with cross-references to main entry	72.8%
Not separate entries, but main entry should include fields for title changes for each edition	8.9
Simply note which language editions exist	15.8
Ignore permutations by differing language editions	1.3

Q: Should geographic editions and special group editions be included if the only change is in advertising content?

Yes	No	Number of Respondents
20.1%	70.9%	164

Although the tabulations show the vote to be overwhelmingly negative, comments indicated that this would be desirable if it could be determined with accuracy. A number of respondents suggested an indicator just to show that this condition exists.

Of the 131 negative responses, 124 answered a question about the inclusion of permutations generated by these editions.

Yes	No	Number of Respondents
40.3%	59.7%	124

Again the comments reflected the feeling that editions should be included especially when the difference is more substantial than advertising content.

Copyright Information

Q: Should an element be included which will denote whether or not material in a serial is copyrighted?

Yes	No	Number of Respondents
50.3%	49.7%	165

Q: If respondent answered "yes", he was asked: When an entire piece is copyrighted, which of the following items might be included?

Choice	Percent Answered Include	Do not Include
Covered by international copyright?	78.5%	21.5%
Country where copyright is held?	70.0	30.0
Name and address of copyright holder?	56.3	44.7
Copyright restrictions, if any?	69.2	31.8

Number of respondents: Approximately 78

Respondents would like to have this information, but they felt that the problems involved would be too great to make inclusion feasible.

Physical Form

Q: Should an element be included to denote the physical form in which an item appears, e.g., printed form, microfiche?

Yes	No	Number of Respondents
95.8%	4.2%	167

The general feeling was that physical form should be noted only when not conventional, especially if special equipment would be required. Many respondents indicated a desire to know if the material was also available in other forms.

Collation Information

Q: Which of the following elements should be included (as originally issued, not as locally bound)?

	Mandatory	Desirable	No Opinion	Not Desirable
Size of serial	21.2%	38.2%	12.1%	28.5%
Illustrations	25.5	39.4	12.7	22.4
Page makeup (e.g.) one or two columns	3.0	9.7	21.8	65.5
Average number of pages per issue	6.1	19.4	17.0	57.6
Average number of pages per year	7.3	24.4	16.5	51.8
Pagination scheme	17.6	30.9	13.9	37.6

Number of Respondents: Approximately 165

Many people said that because of the difficulty of keeping this information up-to-date, as opposed to its value, these items were some of the least desirable in the program.

PART X. DATA ELEMENTS RELATING TO FREQUENCY AND PUBLISHING

Publication Status

Q: Which of the following conditions should be built into a data element relating to publication status?

Publication Status	Should be Included (Percent)
Dead-ceased publication	98.2%
Currently published	97.6
Merged with (a title)	95.2
Supplement to (a title)	92.8
Subseries of (a title)	91.0
Under another title	91.0
Temporarily suspended	86.1

One group of respondents said that more than one of these conditions could occur at once. Some suggested additions to this list such as:

Status unknown

Absorbed

Allocation of space for local status

Establishment of Entries

Q: Should there be an element denoting how an entry has been established (e.g., from one issue, established by L.C.)?

Yes	No	Number of Respondents
57.8%	42.2%	166

Some respondents commented that a uniform method of entry should be used. Others felt that unless this item was included there would be difficulty in communication, particularly in regard to entries contributed to the program by other institutions.

Pattern of Publication

The questions relating to pattern of publication should really not have been included in the questionnaire and as such were somewhat misunderstood. These elements are primarily related to machine functions and the programming for control and update use. However, a tabular summary of the positive responses follows:

Data Element	Yes (Percent)	Mandatory* (Percent)
Dates of Publication and Receipt		
Stated publication date		94.0%
Receipt date		51.2
Actual date of publication		33.1
Numbering Patterns	86.7%	
Former Frequency Data		
Frequency pattern	50.0	
Number schemes	26.3	
All three dates of publishing	11.3	
No former frequency data	46.9	
Bibliographic Volumes per Year, Issues per Volume	85.0	
Indication Suspensions of Publication	93.4	

* Questions were not all phrased for the same type of response.

Reporting Period

Q: Should an element exist giving dates for the reporting period (e.g., for report year ending June 30)?

Yes	No	Number of Respondents
78.0%	22.0%	164

Some respondents felt that a designation was necessary for irregular coverage. Others felt that this information should be included in the descriptive notes, but not be retrievable.

Starting and Ending Dates, First and Last Published Items

Q: For changes in titles, should the program give both of these kinds of information?

Both	66.3%
Starting and ending dates only	26.4%

Number of respondents: 143

The remainder of the 163 persons felt that the program should either include only the first and last published items or include neither kind of information.

PART XI. HOLDINGS, UNION LIST ASPECTS, SUPPLEMENTS, AND TPI

Nonspecific Holdings in the National Record

Q: Do you agree with the approach of indicating nonspecific holdings in the national record and then using specific holdings at the local level?

Yes	Number of Respondents
84.0%	163

The remainder of the 163 persons indicated reservations or suggested alternatives.

Comments ran from the feeling that anything short of specific holdings is a waste of time, to a statement that no holdings should be included at all. Some respondents indicated that the only practical approach was to include detail at local or regional levels only. A few thought that the three national libraries as well as selected regional libraries should list complete holdings. Several respondents indicated a preference for the New Serials Titles form, i.e., noting missing and incomplete volumes.

Q: Is the use of first and last volumes, with appropriate dates, sufficient?

Yes	No	Number of Respondents
78.7%	21.3%	164

Again, many respondents favored the use of the NST form for showing broken files.

Specific Holdings

Part of the question in the questionnaire, relating to nonspecific holdings in the national record, stated . . . "provision would be made to read in [to the Serials Data Program data file] on a local basis, additional specific holdings . . ." The following table summarizes the positive responses to the data elements listed.

Data Element	Percent Positive Responses
Volume Numbers	98.1%
Series Numbers	90.6
Year Used as Volume Number	87.8
Supplement Information	84.9
Issue Number	75.5
Cumulative Indexes (Holdings)	76.3
Special Issue Information	71.2
Part or Section Number	63.5
Separately Published Indexes	59.0
Date of Issue	57.1
Number of Copies Held	27.7
Sample Issues	18.6

Holdings Terms (Volume, Part, etc.)

Q: Should foreign language terms such as band, teil, etc., be used or should they be converted to the English equivalent?

Use English Equivalent	58.2%
Use Foreign Language Terms	41.8%
Number of respondents	165

Those who favored foreign language terms said that a precise English equivalent was not always available.

Geographic Base for Holdings Statement

Respondents were asked how wide a geographic base should be used to select institutions for reporting holdings. The question may have been misunderstood for the response to this question was inconclusive.

Dates With No Western Equivalents

Q: Is it sufficient to translate such dates into the usual notation, the Gregorian calendar?

Yes	No	Number of Respondents
78.5%	21.5%	158

Some respondents stated that for some dates both forms should be used.

Loan and Retention Policy

Q: Which of the following kinds of information related to the retention, reproduction, and loan policies of individual libraries is needed for a serial?

Kind of Information	Positive Response
Will library reproduce?	88.1%
Loan through interlibrary loan networks?	86.1
Loan directly?	74.8
Microforms furnished?	74.2
Charges for any services?	70.9
Retention policy?	59.6
Applications needed?	50.3

Total respondents: 151

Several respondents remarked that this information may change frequently and they questioned whether it could be kept current. Others suggested that these items should be carried in a separate directory, and not in the main data bank. Respondents told the interviewer that these elements should be summarized for each institution, not for each title.

Supplements

Q: Which of the following items should be included in the Serials Data Program?

Type of Information	Mandatory	Desirable	Not Desirable	No Opinion
Supplements: Existence	84.2%	12.7%	0.6%	2.4%
Supplements: Numbering Schemes	72.6	14.6	6.7	6.1
Supplements: Frequency	68.9	20.7	5.5	4.9
Supplements: Titles	67.9	22.2	5.6	4.3
Supplements: Receipt Status	60.1	23.4	7.0	9.5

Number of respondents: I varied from 162 to 165.

Some of respondents commented on the handling of supplements via a separate record, while others expressed the need for handling major supplements either as a monograph or as a separate serial. In general, there was the feeling that this type of information should be carefully considered.

Special Issues, Addenda, Errata

Q: Should the following be handled as supplements or should they be handled separately?

Data Element	Handled as Supplements	Handled Separately	Don't Include As Element
Special Issues	56.9%	35.9%	7.2%
Addenda	28.0	21.0	51.0
Errata	9.0	19.2	71.2

Title Page and Index Information

Q: Which of the following types of information should be included in the Serials Data Program?

Type of Information	Mandatory	Desirable	Not Desirable	No Opinion
How title pages are obtained	51.8%	28.7%	9.1%	10.4%
When separate title pages are published	44.8	32.7	11.5	10.9
How index is obtained	58.8	25.5	8.5	7.3
When separate index is available	59.1	24.4	9.1	7.3
Type of index	33.3	47.9	12.1	6.7
How table of contents obtained	43.9	33.5	12.2	10.4
When table of contents published	48.1	29.6	12.3	9.9

PART XII. OUTPUT

A number of possible products and services of a national serials data program were included in one part of the questionnaire. It should be noted that much of this section was based upon the assumption that everyone would have access to a computer; therefore, the results should be evaluated with this in mind.

The following is a brief discussion of some of the products and services which respondents felt were of top priority.

National Union List

A large proportion (82.2% of 163) of the respondents said that a top priority product of the program was a machine-readable union list, which, if maintained properly, could be used to generate other permutations at the local level, e.g., regional lists. Other comments included the hope that the product would be more comprehensive and more up to date than the present third edition of the National Union List of Serials.

Reports on Serials

Reports on changes in serials, new titles, and ceased titles rated high on the list of desired products, with over 88% considering these items as "top priority."

Universal Numbering System

Again, a system for registering every serial and for assigning it a unique registration number was considered an important service of the Serials Data Program. Such a system was given a "top priority" rating by 56.0% of the 159 respondents who answered the question dealing with this data element.

Traditional Cataloging Forms

There was some confusion over the phase "cataloging forms," that is, respondents asked whether the questions dealt with the physical form of the card catalog or with the format of the main entry record for a title.

Q: Will there be a demand for keeping traditional cataloging forms in the serials program?

Yes Definitely	Yes Not Crucial	No Demand	Number of Respondents
33.8%	43.9%	20.3%	148

Q: Is the cataloging form a problem if the serials program enables the user to retrieve information at any access point?

Yes	No	Number of Respondents
39.9%	60.1%	148

Need for Library of Congress Cards

Q: How much need is there for LC printed cards for serials?

	At Present	When Automated Record Available
Needed	63.3%	21.8%
Needed, but can be done without	23.1	34.6
No need	13.6	43.6

Respondents stated that LC printed cards would be needed as long as card catalogs existed, which is likely to be for many years to come, especially for libraries not having automated records.

I. ALTERNATIVES

A. Content

A proposed communications format for serial publications may be found in Appendix II of this report. Although some of the data elements which were identified in Task A of Phase I of the National Serials Data Program have been eliminated from this format, the document does present those data elements seemingly required for an optimum utility. It should be noted that the costs which were calculated for such a system did not take into account the cataloging activities and other intellectual efforts necessary to provided the information.

1. Alternatives for the Format

The data elements in the variable fields of the format have been divided into four categories according to the following criteria: 1) those essential for basic identification and location; 2) those needed for reference purposes primarily to serve reader requests; 3) those used to provide more precise identification for bibliographic purposes; and 4) those elements which might be needed for more precise identification of a serial but could be deleted with minimal loss.

Each of the following categories should be viewed as modules of an automated serials data bank. Category 1 contains only the data elements necessary for a minimal system on the national level. Category 2 added to Category 1 would enable library personnel to provide reference service to users. Category 3 added to Categories 2 and 1 would give more precise identification of a serial, and so forth. The costs per title have been cumulated accordingly, and these figures reflect only the costs for current serials. Cost data may be found in Table I. of "III. COST ASSUMPTIONS".

*Indicates that representation not decided upon or data not available in the test sample from which cost figures were derived.

Category 1 - Basic Identification and Location

*022 Standard serial code
100, 110, 111 } Main entry
245 Full title
260 Imprint
362 First published item and starting date
780 Preceding entries
785 Succeeding entries
*800 Non-specific holdings
800 Specific holdings (Not in format)

Category 2 - Needed for Reference Purposes

041 Language of text and abstracts[1]
050 LC Call Number
051 LC copy statement
082 DDC number
140 Entry as it appears on the piece
*145 Abbreviated entry
247 Former titles and/or title variations
310 Frequency
362 Last published item and ending date
362 Suspension of publication
368 Items not published
370 Irregularities and peculiarities in numbering
510 Indexing and abstracting coverage
710 Issuing Bodies
711 Issuing bodies - conferences
760 Main series entries
762 Subseries entries
770 Supplement/special issues entries
772 Parent record entries

Category 3 - Needed for More Precise Identification

040 Input/cataloging source
041 Language from which translation made[1]
*060 NLM Call Number
*061 NLM Copy statement
*070 NAL call number
*071 NAL copy statement
086 U.S. Supt. of Docs. number
200 Varying forms of title
205 Distinctive title
208 Added title page title
246 Titles in other languages
250 Edition statement
*320 Frequency control
*321 Former frequency control
*330 Publication pattern
*331 Former publication pattern
371 Cumulative indexes published
*500 General notes
*540 At head of title note

---

[1]Costs for this tag were divided into proportional parts since the cost of the various elements in the tag were not calculated separately.

600
1 Subject headings[2]
652
700 Personal names associated with serial
712 Former issuing bodies
714 Sponsoring agency of conference
730 Uniform title heading
765 Original entries for translations
*767 Entries of available translations
*777 Issued with entries

Category 4 - Minimal Loss

010 LC card number[3]
015 National bibliography number[3]
025 Overseas acquisitions number[3]
*035 Local system number
*030 Coden
041 Language of available translations
*072 NAL subject category number
*080 UDC number
130 Uniform title heading (Not in format)
240 Uniform title (Not in format)
241 Romanized title (Not in format)
*265 Publisher's address
*300 Collation
775 Other available editions
261 Former places of publication (Not in format)
350 Subscription price
362 Report year coverage
530 Additional physical forms available
506 Limited use note
525 Supplement note
718 Indication that serial is an official organ
*750 Names not capable of authorship

The data elements in the fixed fields and the fields needed for internal machine processing (such as the subrecord directory, subrecord relationship, etc.) were not included in the above categories since these would vary according to requirements of individual serials control systems. Although differing opinions exist on the inclusion of certain data elements under a particular category, it would still be possible to establish priorities by considering the costs of the individual data elements themselves.

---

[2]Subject headings have been placed in this category because they are generally so broad for serials that their use would be limited (as opposed to monographs).

[3]These control numbers would not be needed once the standard serial code is developed.

Alternatives could also be stated for the inclusion or exclusion of a data element within any of the categories. The most important would be in Category 1 where the following courses of action could be implemented regarding holdings:

1. NSDP could include specific holdings of the three national libraries and the non-specific holdings of all other libraries.

2. NSDP could give only non-specific holdings, as is done in the Union List of Serials, for all libraries.

Given the comparatively high cost of including specific holdings (approximately one-third of the total cost of Category 1), the national system may have to settle for only non-specific holdings representation.

B. Scope of Program

The optimum serials data bank would include all serial publications that had ever existed with full bibliographic information pertaining to each one. Taking into account the alternatives to the format and file structure, the Serials Data Program will still be forced to limit the number of records in the system while attempting to provide maximum service to its users. One decision that must be made is whether the program will contain records for only currently published material. The following is a list of possible alternatives on this matter:

Not under consideration

1) All serials, live and dead, would be available in the program on-line (economically unfeasible).

2) Only new serials published after a certain date would be input and available (unsatisfactory since the machine file would control only a very small percentage of serials).

Under consideration

3) All serials currently being published would be input initially on an on-line basis.

Cost per live title

Cat. 1	$72.64
Cat. 2	29.10
Cat. 3	23.20
Cat. 4	16.49
TOTAL	$141.43

4) Only those serials which are currently being published would be placed in the data bank on an on-line basis; any serials which cease publication after time of initial input would be placed in secondary storage.

Cost per live and dead title

Cat. 1	$48.92
Cat. 2	24.03
Cat. 3	16.05
Cat. 4	11.50
TOTAL	$100.50

Since maintenance costs will be virtually non-existent for dead serials, the average cost per title for a data bank containing live and dead serials is less than that for live serials. The total cost for the more inclusive data bank would obviously be higher.

Detailed cost figures for both current and dead serials can be found in Table I of Section II of this report.

The problems associated with the third and fourth alternatives deal with the capability of cross referencing preceding and succeeding entries over a period of time if dead serials are not included in the data bank.

Although suggestions have been made to limit the scope of the Serials Data Program by subject coverage (e.g. science and technology titles), this alternative would not serve the library or library - related organizations to their satisfaction. It would be extremely difficult to define subject coverage of serials when the contents of any given publication can fall into different disciplines (related or unrelated), and such overlapping would not be an uncommon occurrence. In addition, it would not be possible to foresee any additional subjects or changes in subject coverage after initial input into the data bank(s). Also, subject limitations would not allow the various library systems to handle the broad spectrum of serial material which researchers require in such fields as the humanities, arts, or social sciences.

Form of material could be used as another criterion by which to limit the scope of NSDP; however, this alternative was not retained since the basic concept of MARC (Machine-Readable Cataloging) was to provide a communications format flexible enough to handle different forms of material. Hence, records for serials published on microfilm may be input with a minimum number of adaptations.

Another alternative to the scope of the program is to limit the number of serials based on substantive content. Although the original concept

of the program called for coverage of all known serials, it was recognized that some exclusion by substantive content must be accepted particularly on the local level. It would be extremely difficult to establish guidelines for a national system since what may be ephemera to some users may be research material to others. Existing publications such as the Union List of Serials exclude the following types of publications: Government publications (except periodicals and monographic series); United Nations publications; administrative reports of societies, universities, corporations, etc.; almanacs; gift books; American newspapers; English and other foreign newspapers after 1820; law reports and digests; publications of agricultural and other experiment stations; local religious, labor, and fraternal organizations; boards of trade; chambers of commerce; national and international conferences and congresses, etc.; house organs (unless of technical and scientific value); alumni, undergraduate, and intercollegiate fraternal publications; trench papers; and in general all titles having a highly limited or ephemeral value. On the other hand, New Serial Titles excludes only newspapers, looseleaf publications, books-in-part, motion pictures, filmstrips, and phonorecords. The Editor of New Serial Titles has also indicated that "ephemera" were excluded if personnel at LC could make this decision by looking at the piece itself.

In most instances, the Union List of Serials has attempted to define substantive content by stating the types of publications to be excluded, e.g., publications of local religious, labor, and fraternal organizations. The NSDP user survey had attempted to determine what might be excluded from the program through its questions on the types of publications; however, the results were inconclusive except that there was general dissatisfaction with the definitions of the "types." Unless definitive guidelines can be developed for the program, it is likely that subjective judgments as presently being applied for New Serial Titles will need to be continued.

The National Serials Data Program could also limit its coverage by language, e.g. including only those serials in Roman alphabets. This alternative, however, was not retained in view of the overwhelming opinion of respondents in the user survey and the three national libraries that it would be preferable to input such material with information in romanized form than not to include them at all.

C. File Structure

1. Storage Sites

Several possibilities exist for determining the number of storage sites which should be included in the design considerations of the National Serials Data Program. The possibility was explored of planning a system with a large number of stores (up to 1,000) maintained from a central point with complete or incomplete data banks to serve regional users. This plan was ruled out because of the problems of financing and controlling such a network. The alternatives retained for consideration are:

1. One central site
2. Three sites (LC, NLM, NAL)
3. Approximately ten sites (One central site plus nine regional sites)*

In order to bring a National Serials Data Program into being, some organization or organizations must assume or be assigned the responsibility for operating or funding the effort. These organizations will provide the fundamental guidelines within which a national program must operate. A number of "environmental related" factors should be considered in conjunction with the assignment of responsibilities of a national program and the determination of its location(s).

1. The completeness and convenience of the physical access to the ... serial literature that is provided.
2. The availability of, and access to, a supply of skilled bibliographic analysts capable of handling the widest range of foreign languages.
3. The compatibility of operational functions and congruence of program interests between the Serials Data Program and its parent organization.
4. The availability of means for continued program funding and effort.**

2. Linked Files

Closely related to any considerations of the number of storage sites are the decisions regarding the number of data banks to be included in the system, the contents of the data banks, and the method of linking between the existing banks.

The alternatives for the number of data banks are as follows:

1. One site
   a. One central data bank with local linkages.
   b. One central data bank which is composed of separate files consisting of categories of data as described under Alternatives for the Format. These files would be linked locally to form one data bank.

---

*For the purposes of this report, the nine regions by which the U.S. Office of Education, Division of Library Services and Educational Facilities, has divided its organizational coverage have been used.

**Information Dynamics Corporation. A serials data program for science and technology. Prepared by William A. Creagar and David E. Sparks. Reading, Mass., 1965. p. 131.

2. Three sites
   a. One central data bank with three national linkages.
   b. Three independent data banks with local linkages to each bank (with delegated subject coverage).
   c. Categorized files comprising one data bank linked to the three national libraries (i.e. four categories of data spread among the three national libraries.)

3. Approximately ten sites
   a. One central data bank with regional linkages.
   b. Approximately ten independent data banks with regional links to each bank.
   c. Categorized files comprising one or more data banks linked to the regional libraries.

The alternatives for the record content of the data bank are as follows:

1. All records duplicated at each site.
2. Records found at any site on a selective basis (depending on any criteria established) with the central site containing all entries.
3. Records found at any site on a selective basis with files pulled as necessary.

Regardless of which alternative is decided upon, it is imperative that linkage, input, and update methods be consistent with the national system and not be attuned to local machine manipulations. Local products may be generated based on local needs. Similarly retrieval methods will reflect local requirements. Uniform standards are essential to safeguard against the possibility of file destruction at the national level. All of these alternatives would entail enormous linkage and updating routines, and the basic question seems to be whether the initial responsibility for updating should rest entirely with the central site or with regional sites who would transmit data for selective titles to the central site (or actually, there might be a network of reporting libraries as presently found for the New Serial Titles project). In either case, the central site would still have the responsibility of maintaining the central file and possibly transmitting update information at regular intervals on a subscription basis.

The organization of the files for the Serials Data Program continues to be a major problem. Two methods (concerning entry) exist by which records could be organized:

1. Including all information about the serial under the latest title and/or entry, with links from earlier titles/entries to the latest.

2. Setting up new records for changes in title and/or entry with appropriate links to predecessors and successors.

By keeping all bibliographic and holdings data on a single record under the latest title/entry, total storage space and printing costs would be reduced when producing a union list of serials or any other print-out. On the other hand, a method of using split entries would keep the size of individual records to a minimum, would allow reformatting to the first method from the second since it is easier to pull together split entries than to do the reverse, and would allow for faster retrieval because of easier sequencing methods. It has been recommended that the second approach be applied to the Serials Data Program although there were still many reservations.

Although the file structure of the serials system at the central site or regional sites would be primarily a local design problem, this would have national implications in terms of serving users of the national system efficiently and rapidly. Parallel files within any system could be established so that truncated records would be available for functions such as accessioning on an on-line basis with rapid access. Other functions such as binding or acquisitions could be performed satisfactorily with less rapid access of data. Reference or bibliographic functions would be more difficult to segregate by required access time because of varying circumstances under which the information is needed.

II. POSSIBLE COURSES OF ACTION

Differing opinions exist as to how the Serials Data Program should implement its automated serials control system. For example, there are advocates for both an intensive and extensive approach. One could limit the amount of bibliographic information associated with each record so that a larger number of titles could be included in the system, or the number of titles included could be limited while full bibliographic information is represented for each record.

There are also options to begin with either a pilot project or the full system. Even the pilot would necessitate the existence of a fully automated serial control system at LC or any other central site in order to provide continuous update capabilities. The question of which body of serial material to use for the pilot project has been approached in two ways: using new serials entering the system or taking a separate sub-file of a main serial record. The latter method has the advantage of being able to test programs on existing and more complete files while the former enables the designers to proceed without interfering with the day-to-day manual operations. Implementation of the full system initially would probably require a fairly small data bank or a reduction in the amount of bibliographic information carried.

Another alternative is to use the existing files of other organizations that have already been converted to machine-readable form to save on the costs of conversion in the national system. Organizations in both the public and/or private sector would have to volunteer the use of their files for this purpose.

Planning for the national system could not proceed beyond a general outline nor could costing have meaning until a format for serials was quite highly developed. The format is also needed by other library centers which are planning automation. These were the overriding reasons for going forward with development of the format for serials. This development produced a preliminary format to be used as a planning tool for the system design and as a basic standard for transmittal of data.

Effort can now be constructively expended toward building a data base which will serve as the foundation for a national serials system. This effort could be confined to the data elements included in Category 1. The reasons for this limitation are persuasive even though it is recognized that, in principle, costs are increased each time the same items must be handled for inclusion in a data bank. This additional cost is comparatively greater for monographs than it is for serials

because each serial record is, in effect, self-purging because the information about a serial must be updated and the record must of necessity be addressed, the item re-examined, and the details of existing elements changed.

An initial National Serials Data Program should include only the data elements included in the first category for three reasons. First, they are the elements most readily agreed to for inclusion in the system. The user survey and other consultations have shown general agreement in regard to the necessity of this basic cluster. Secondly, this cluster contains those elements which are essential to the user of a larger expanded set of elements. It will be remembered that benefits are obtained from the cumulative availability of data elements. That is, basic benefits are obtained from Category 1, additional benefits from Categories 1 and 2, and still others from Categories 1, 2 and 3. Thirdly, basic system design indicated that each record would probably best be maintained in memory not as a monolithic whole but would rather exist as sub-records in as many as three linked files. Economy seems to dictate that elements in the third category can be held in memory which is less immediately available and is therefore less expensive per element stored.

The options can most readily be explained by means of two simple systems. Suppose that each record R_i is made up of three element clusters, $E_{i,1}$, $E_{i,2}$, and $E_{i,3}$. The design could create a single file which would be made up of records $R_1 \ldots . R_n$ and each R including $E_{i,1}$, $E_{i,2}$ and $E_{i,3}$. The file would be stored as:

R_1 ($E_{1,1}$, $E_{1,2}$, $E_{1,3}$)

R_2 ($E_{2,1}$, $E_{2,2}$, $E_{2,3}$)

.

.

.

R_n ($E_{n,1}$, $E_{n,2}$, $E_{n,3}$)

Suppose that when the file is addressed, only $E_{i,1}$ is desired from all records. A query is made to scan each R and to select from the file each $E_{i,1}$ to output ($E_{1,1}$, $E_{2,1}$ ... $E_{n,1}$). If the entire content were desired, the query would designate $E_{i,1}$, $E_{i,2}$ and $E_{i,3}$ to output the entire file. Elements of classes 1, 2 or 3 are, in effect, segregated at the time of output, and a report is created by consolidation of data from all R's.

The second option is, of course, to set up three files designated E_1, E_2, and E_3 as follows:

File	E_1	E_2	E_3
	R_1	R_1	R_1
	R_2	R_2	R_2
	.	.	.
	.	.	.
	R_n	R_n	R_n

If only $E_{i,1}$ were desired, file E_1 would be queried. If all information were desired, all files would be queried, and at time of output, $E_{1,1}$, $E_{1,2}$, and $E_{1,3}$ would have to be related to produce R_1.

Economics would dictate which is the better design for a given situation. Because of the rather compelling reasons given, the second design option would seem to be superior with the provision that when a full system is implemented, the file might be reconstructed to take advantage of new machine developments in storage.

The difference between the two approaches exists essentially within the machine and should not require differences in procedure at input and output. A record representing a title would be tagged and delivered to the machine in both cases. Normal procedure would place all such records in a single file under the first organization. The second procedure would divide each record into components as specified (e.g. into Categories) after the record had been delivered to the machine. Similarly at output, if a complete record were called up by a query statement, it would address the appropriate file, find the record designators (linking terms), and create a complete record from the cluster of elements in each file. It follows that processing costs for record analysis and synthesis would be increased by the split approach at input and output, but the reduction in searching costs would offset these increases for certain configurations of memory and processing components.

The cost studies have supported opinions generally held and have developed some new and somewhat surprising concepts. It is quite true that the cost of a system increases as the number of data elements per title increases. It is not true that the increase is linear either in regard to elements or in regard only to number of characters cumulatively included in the record. Other factors such as memory requirements and frequency of change tend to distort the picture.

III. COST ASSUMPTIONS

The following assumptions were made when calculating the costs for the National Serials Data Program:

1) Since the benefits of the serials system are measured with respect to the potential users or group of users, the costs of the system should be estimated from their point of view. This means that the initial investment (input) costs and annual recurring costs consisting of direct operating costs and maintenance costs for data elements should be considered explicitly.

2) Calculations were made for a time period of ten years.

3) The problem of joint costs arises here with the initial investment since the long-run average initial cost of a data element depends upon what data elements are procurred at the same time. The initial investment costs have been estimated from the data provided by Project MARC and the Technical Information Project at M.I.T. In the MARC Pilot Project Final Report, it was reported that by March 1968, the cost per record was approximately $1.31, and preliminary estimates indicated an average of 500 characters were punched for each title. This implies a cost per character, i.e., long-run average cost per character of $0.0026.

4) The direct operating cost is that of storing information on tape drive or magnetic drum for a particular data element. This is assumed to be expressible in terms of a cost per bit year and the average number of bits in that data element. M.I.T.'s Technical Information Project has provided an estimate based on their serials data program for the cost per bit year, i.e., with eight bits per character, the cost per bit year is $0.00045.

5) The annual maintenance costs for a particular data element depends upon the frequency with which individual records must be changed and updated. As a result, the estimate of the cost of file maintenance depended heavily upon the assumption about the relative frequency of changes in data elements.

6) The total number of currently published serials initially included in the projected data bank is 225,000. This figure was based upon the holdings of the Library of Congress, the National Library of Medicine, and the National Agricultural Library. There are approximately 640,000 distinct entries in LC's combined files (i.e., the 3 x 5 file and the visible file); approximately 390,000 are classified and 250,000 are not classified. An estimated number of 185,000 distinct live titles are contained in the serial record after factoring out all cross-references and dead titles. There are approximately 25,000 titles in NLM and 25,000 titles in NAL that will be included in the data bank. Of these titles, no duplications could be factored out which might be contained in the Library of Congress collection. Thus, the data bank for purposes of determining the cost figures was established to be approximately 225,000 live serials.

7) New serials to be added per year to the system are estimated to be approximately 15,000. This figure is derived from the total number of serials received each year in the three national libraries. The Library of Congress New Serials Titles Section receives reports for approximately 12,000 new titles per year; NLM receives around 1,500 new titles per year; and NAL receives around 1,500 new titles per year. Since an estimated 7,500 titles cease publication each year, the net number of serials to be added to the system was fixed at 7,500.

8) The terms "Category 1," "Category 2," etc., refer to the division of the data elements into four categories as described in II. ALTERNATIVES.

9) All costs are estimated in current dollars and at current equipment purchase prices. No attempt has been made to increase costs to account for increasing labor costs or to decrease costs to account for lower equipment costs per bit.

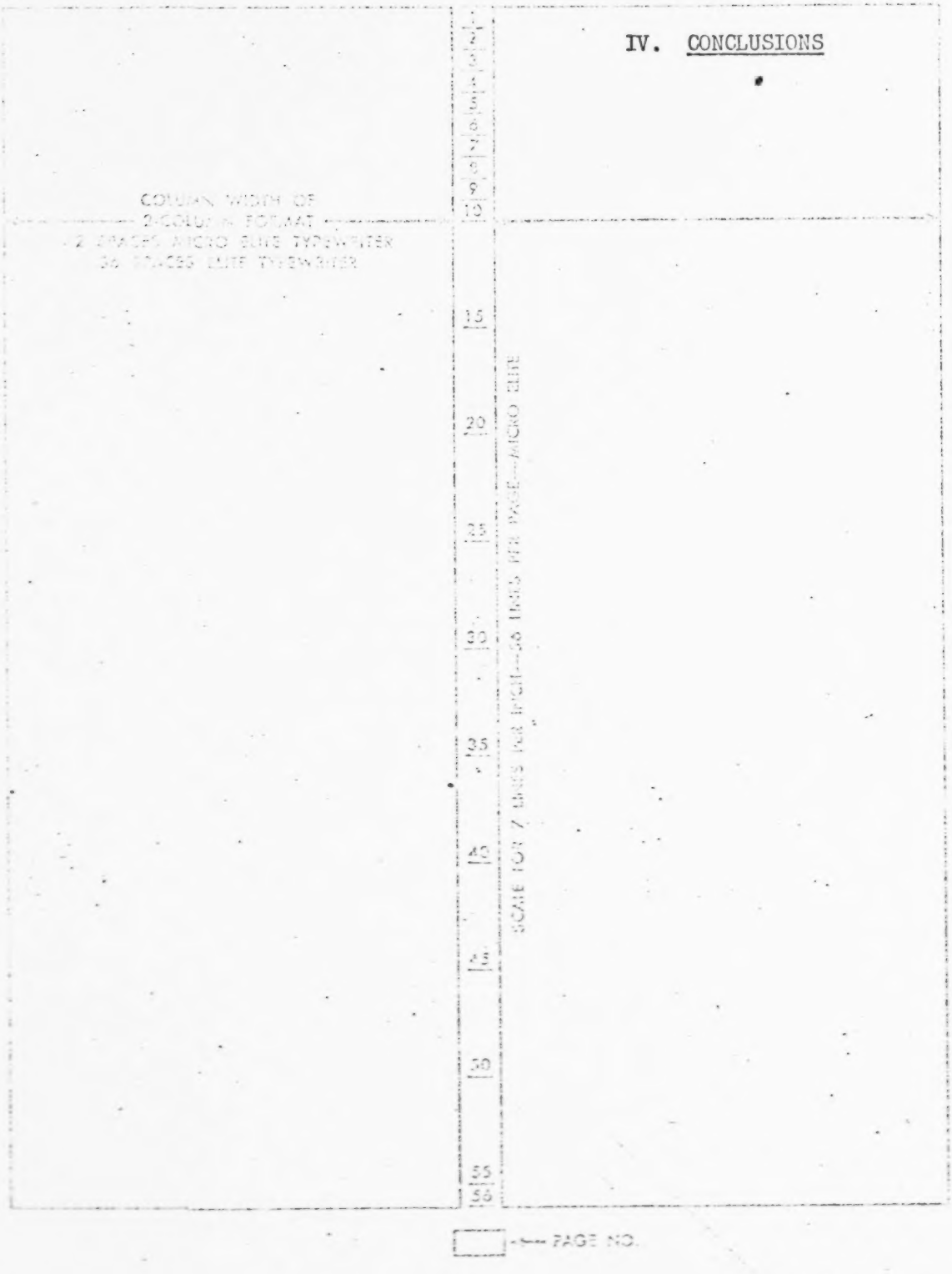
TABLE I. ESTIMATED COSTS FOR NSDP FOR TEN-YEAR PERIOD

	LIVE TITLES	DEAD AND LIVE TITLES
Cost Per Title		
Category 1	$72.64	$48.92
Category 2	29.10	24.03
Category 3	23.20	16.05
Category 4	16.49	11.50
Cumulative Cost Per Title		
Category 1	$72.64	$48.92
Category 2	101.74	72.95
Category 3	124.94	89.00
Category 4	141.43	100.50
Total Input Cost Per Category		
Category 1	$111,969	$201,781
Category 2	55,414	100,933
Category 3	45,416	84,880
Category 4	41,228	58,139
Total Storage Cost Per Category		
Category 1	$135,916	$284,781
Category 2	67,240	142,291
Category 3	55,226	119,061
Category 4	23,658	58,139

LIVE TITLES		DEAD AND LIVE TITLES
Total Maintenance Cost Per Category		
Category 1	$12,542,733	$19,076,400
Category 2	3,181,406	9,489,600
Category 3	3,843,168	6,690,960
Category 4	3,550,437	5,418,000
Total Cost Per Category		
Category 1		
Input	$ 111,969	$ 201,781
Storage	135,916	284,236
Maintenance	12,542,733	19,076,400
Total	$12,790,618	$19,562,417
Category 2		
Input	$ 55,414	$ 100,933
Storage	67,240	142,291
Maintenance	3,181,406	9,489,600
Total	$3,304,060	$ 9,732,824
Category 3		
Input	$ 45,416	$ 84,228
Storage	55,226	119,061
Maintenance	3,843,168	6,690,960
Total	$3,943,810	$ 6,894,901
Category 4		
Input	$ 19,489	$ 41,228
Storage	23,658	58,139
Maintenance	3,550,437	5,418,000
Total	$3,593,584	$ 5,517,367

LIVE TITLES		DEAD AND LIVE TITLES
Cumulative Total Cost Per Category		
Category 1	$12,790,618	$19,562,417
Category 2	16,094,678	29,295,241
Category 3	20,038,488	36,190,142
Category 4	23,632,072	41,707,509

IV. CONCLUSIONS


COLUMN WIDTH OF
2-COLUMN FORMAT
2 SPACES MICRO ELITE TYPEWRITER
96 CHARS ELITE TYPEWRITER
SCALE FOR 7 LINES PER INCH—66 LINES PER PAGE—MICRO ELITE
PAGE NO.

INTRODUCTION

The primary objective of the National Serials Data Program is the design and implementation of a national serials information system. The goals of this system are the provisions of timely, accurate, and comprehensive information about serial publications within a framework of quantitative efficiency and agreed upon cost effectiveness. The techniques for initiating these goals include: adherence by libraries and information centers to standards for machine representation and transmission of information about serial publications; the creation of a national network which encompasses and builds upon existing serial information systems; the delegation to appropriate national, regional, and local agencies responsibility for capturing as close to the generating source as possible serials' information in machine-readable form; and the structuring of a workable method for assigning costs on the basis of use, and crediting costs on the basis of services which have been provided the national serials information system by member libraries and information centers. Critical to the long range success of the national serials data program is the delegation to the three national libraries of responsibility for augmenting and integrating their present serials information dissemination programs into a workable national serials information system. Critical to the short range success of the National Serials Data Program is the adherence to standards being proposed as the format for machine-readable cataloging records of serial publications. Both long-range and short-range plans presuppose an evolving network of users and information suppliers. The provision of a machine-readable format in the phase-one report of the national serials data program thus serves as the first step toward the creation of both a short range, evolving system and a long range system with responsibility for designated national, regional, and local serials information processing centers emerging through the exchange of serials information. Both short and long plans require further systems design and additional administrative decisions. These must be formulated in light of alternatives identified during the first stages of the National Serials Data Program.

The alternatives presented in this report were discussed with the three national libraries, with the Joint Committee on the Union List of Serials and with a special committee of the National Federation of Science Abstracting and Indexing Services. These alternatives, summarized in the table which follows, (Figure 1) serve as the basis for the conclusions drawn from the Phase I study. Decisions should be based upon the means of collecting possible data elements, reactions from potential users as expressed in the user survey and in conferences, and the results of the cost survey.

Decisions must be reached upon the composite set of alternatives since all of the alternatives will have interdependent effects on the design of an eventual national serials system. Because of this interdependence,

conclusions are not presented as decisions made with respect to individual alternatives, but as overall conclusions made in regard to the Phase I effort.

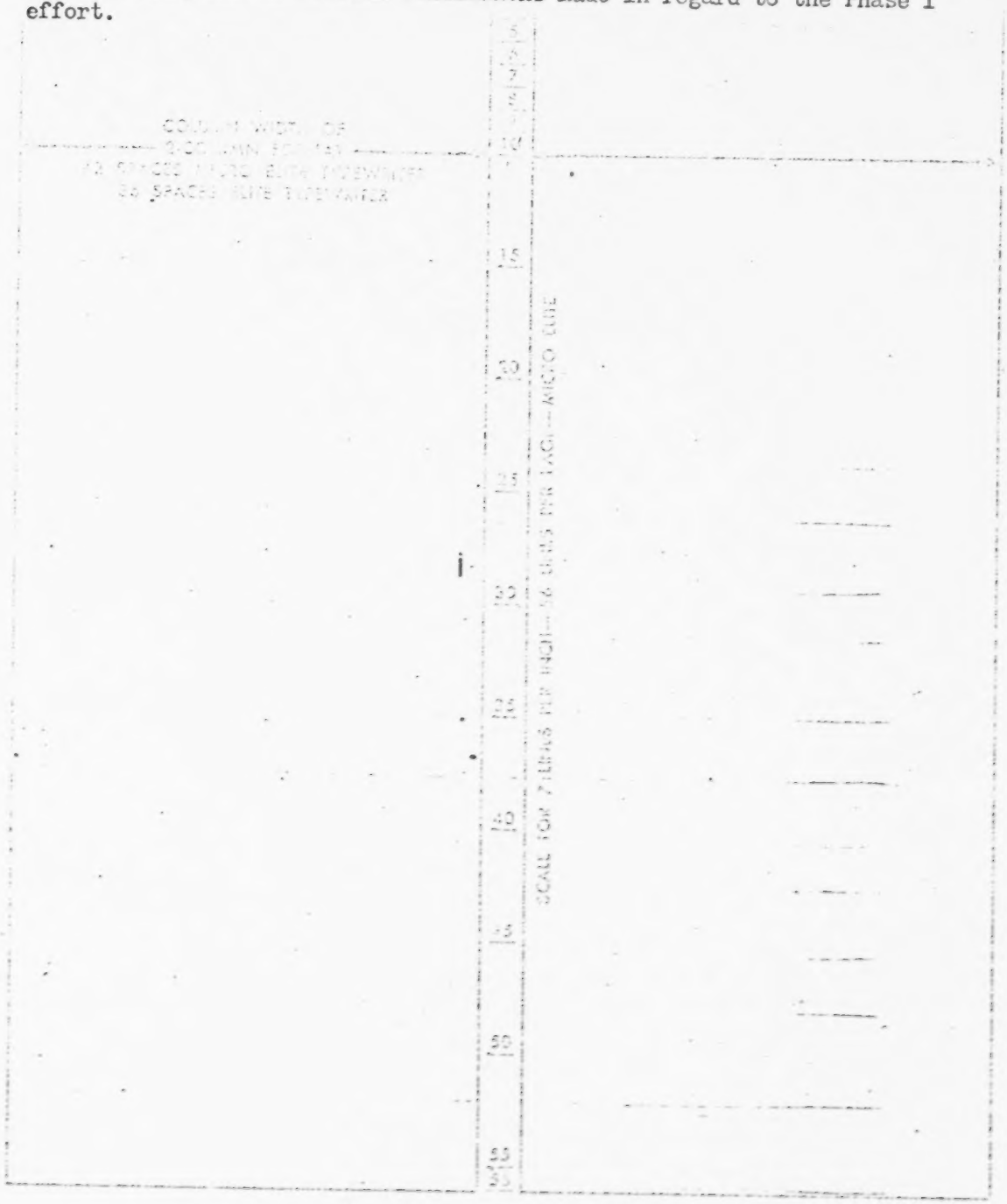


Figure 1

SUMMARY OF ALTERNATIVES

A. Objectives of the Format:

1. Basic identification and location
2. Meeting reference needs and user requests
3. Identification sufficiently precise for definitive bibliographies
4. Highly precise description with additional access

B. Scope of Coverage:

1. Vital Status

   a. All serials alive and dead
   b. Only new serials published after a certain date
   c. All currently published serials
   d. All currently published serials with cessations in secondary storage

2. Subject

   a. All subjects
   b. Limited to a subject category, e.g., science

3. Substantive Nature of Content

   a. All known serials, including ephemera
   b. New Serial Titles definition

4. Types of serial publications

   a. Serials defined in broad context
   b. Only periodicals
   c. New Serial Titles definition

5. Language

   a. All languages
   b. Only English
   c. Other subdivisions

C. File Structure

1. Storage sites

   a. One central site
   b. Three sites (Three National Libraries)

c. Limited number of regional sites maintained from one central site
d. Large number of primary sites maintained from one central site

2. Linked files (Number of data banks)

One site:

a. One central data bank with local linkages
b. One central data bank composed of categorized files

Three sites:

c. One central data bank with three national linkages
d. Three independent data banks with local linkage to each (delegated subject coverage)
e. Categorized files comprising one data bank linked to the three national libraries

Limited Number of regional sites:

f. One data bank with regional links
g. All independent data banks with regional links
h. Categorized files comprising one or more data banks linked to central

3. Record Content of Data Bank

a. Central including all entries; other sites selective
b. All sites selective
c. All records duplicated at all sites

4. Record Organization

a. Successive entries (AA)
b. All information under latest entry (ALA)

D. Implementation

1. Approach

a. Extensive - many titles with limited bibliographic information
b. Intensive - limited titles with full bibliographic information

2. Pilot project

a. New serials entering the system
b. Separate subset of current serials
c. Existing machine-readable files of other organizations

Format for Machine-Readable Records

The exchange and adaptation for local use of information about serial publications requires adherence to standards. The extent and exactitude of this adherence varies directly with the degree of mechanization required. Manual systems, for example, which control small collections of serial publications, and which do not demand instantaneous access do not require the extensive and always explicit adherence to standards which characterize mechanical and electronic serials information centers. Systems which control large serials collections and which have requirements for quick responses to requests by users, on the other hand, must adhere to extensive standards, both for entering their own records in their local serials information system and for adapting records which have been machine input by others and transmitted as machine-readable data or in machine-readable form. Once developmental and capital costs have been absorbed, automated serials information systems will include in their total system at lower operating costs information about serials which is presently disseminated in non-machine readable form. More importantly, this information will be disseminated on a timely basis; secondly, once input and corrected it will be both accurate and comprehensive; and thirdly, it will be an integral part of a larger, augmented serials information system.

Adherence to the proposed format is thus the first step; agreement on a standard which uniquely identifies in a coded form each serial publication is a second step; and the structuring of a network for exchanging machine-readable information about serial publications is the third stage in the creation of a viable national serials information system.

Further definition of the format emphasizes that the most important initial contribution that the Serials Data Program should make is the development of a standard format for serials. The development of a standard format could then form the basis for a national serials system. If standards are set initially, a national serials data system will to some degree evolve naturally.

The data elements defined in the preceding section (III. Alternatives, p. 1) are divided into four categories in terms of use. Each successive category describes the serials with increasing precision and expanding comprehensiveness and allows the user to develop tools and receive products and services from the machine record to accomplish both ends.

In determining the inclusion of the data elements from these four categories, conclusions were based on several interdependent factors: 1) the products and services required from the system as reflected from the user survey and other conferences; 2) the uses to which these products and services would be put; 3) the data collection methods necessary to acquire any given

data element or set of data elements; and 4) the fact that products and services from the system must update, augment, or replace existing products and services.

The user survey response showed that highest priority should be given to producing reports on new titles (92.6%), reports on changes in title and cessations (91.9%), and the compilation of a national union list (82%). Cataloging data for current serials were also given high priority (84.1%) as a needed service that the above mentioned products should provide. The uses of these products and services by categories of users and the categories of data elements needed to fulfill these uses are outlined in Chart IC.

To fulfill the various needs of the potential users, no one single data collection method will suffice (see Chart IA for data collection methods). A combination of several or all of the collection methods must be used. If, for example, the data elements were to be limited to only those in Category I, more than one data collection method would still need to be employed.

Presently existing printed tools such as the New Serial Titles, and the Library of Congress printed card have wide acceptance in the library and library-related communities. Implicit in the response from the user survey and the needs expressed through such surveys as the "Kulhman Survey" is the conclusion that the Serials Data Program should provide information at least to the level of detail as found in New Serial Titles.

The conclusions are such that the format must be developed to fulfill the most important uses. Although it could perhaps be argued that the format should contain only those data elements from Category I (as these elements are requisite for all users and essential to those users that require a larger number of data elements) there are many factors that negate any decision to use only those elements from Category I.

1. Chart IC shows that the uses to which the products would be put by a majority of users would not be adequately served by the data elements in Category I.

2. Although an individual user could augment his local record to include desired items, costs are greatly increased each time an item is handled.

3. Since no one data collection method will suffice, costs would be substantially reduced if data to fulfill the largest number of high priority products and uses are input once at a single place.

4. Since the central agency must have national control information needed for serials processing at any level, these data elements should be provided once, relieving libraries of this burden and associated costs.

5. The data elements in Category I are not adequate to replace such widely accepted existing tools such as New Serial Titles.

A format (Appendix II) has been developed which implicitly fills the needs of the most important uses. Importance is reflected by responses to the user survey. Implicit use is expressed in terms of data elements required to produce New Serial Titles and the bibliographic record as expressed by the Library of Congress printed card. Because the needs of potential users are not universally consistent, a format is recommended to fulfill the requirements for the highest priority products by focusing upon these two widely used products.

The most productive discussions resulted from considerations of these specific products. Interaction with other libraries and library-related organizations indicated that widespread and direct benefits would result from the ability to issue these products more quickly and with more up-to-date information. It was pointed out that these benefits could accrue not only to libraries which were automating but also to those which intended to continue serial handling and publication of holdings lists on a completely manual basis.

Chart IC shows how the objectives of the format can be achieved in order to provide for identified users. Chart IB indicates that the objectives can be met only as a result of a process of bibliographic verification. This verification operates upon the raw inputs derived from the sources shown in Chart IA.

Verification is a very complex process. It must exist at any level where serials data enter the system from a number of disparate and interested sources. In order that the record be as current as possible, inputs must be normalized as well as possible through the steps of processing.

I-A INPUTS

SOURCES	NEW TITLES	CHANGES	HOLDINGS	A&I
PHYSICAL EXAMINATION				
Deposit, e.g. Copyright	x	x		
Cooperative Acquisitions, e.g. Shared Cataloging	x	x		
Gift-Exchange Agreements	x	x		
Purchase	x	x		
PUBLISHED TOOLS				
Union Lists	x		x	
Regional Union Lists	x		x	
National Bibliographies	x	x		
Periodical and Serial Directories	x	x		
Specific Subject (or Area) Bibliographies	x	x		
Bibliographies of Serials	x	x		
Compilations by type of material, e.g. World List of Future International Meetings	x	x		
PUBLISHED ANNOUNCEMENT MEDIA				
Special Announcements	x	x		
Publisher's Catalogs	x	x		
Library and Publishing Periodicals	x	x		
Distributors Catalogs/Subscription Agent Lists	x	x		
Library Acquisition Lists	x	x	x	
UNPUBLISHED BIBLIOGRAPHIC COMPILATIONS				
A&I Services Lists of Coverage	x			x
Library Card Catalogs	x	x		
Library Serial Records	x	x	x	
ROUTINE INQUIRIES (OR VOLUNTARY COOPERATION WITH)				
Publishers.	x	x		
A&I Services	x	x		x
Professional Societies and Associations	x	x		
Dealers and Distributors	x	x		
Domestic and Foreign Libraries	x	x	x	
SPECIFIC REQUEST FOR ITEM(S)				
by User	x	x		
Recommending Officer	x	x		
Acquisitions Orders	x	x		
Interlibrary Loan Analysis	x	x	x	
ITEM MONITORING				
via Check-in	x	x	x	
Acquisitions	x	x		
Claims	x	x		
OTHER ESTABLISHED MACHINE READABLE DATA BASES	x		x	
Publishers	x			
Librarys and Information Centers	x	x	x	
A&I services	x			x
Dealers	x			

CHART IB

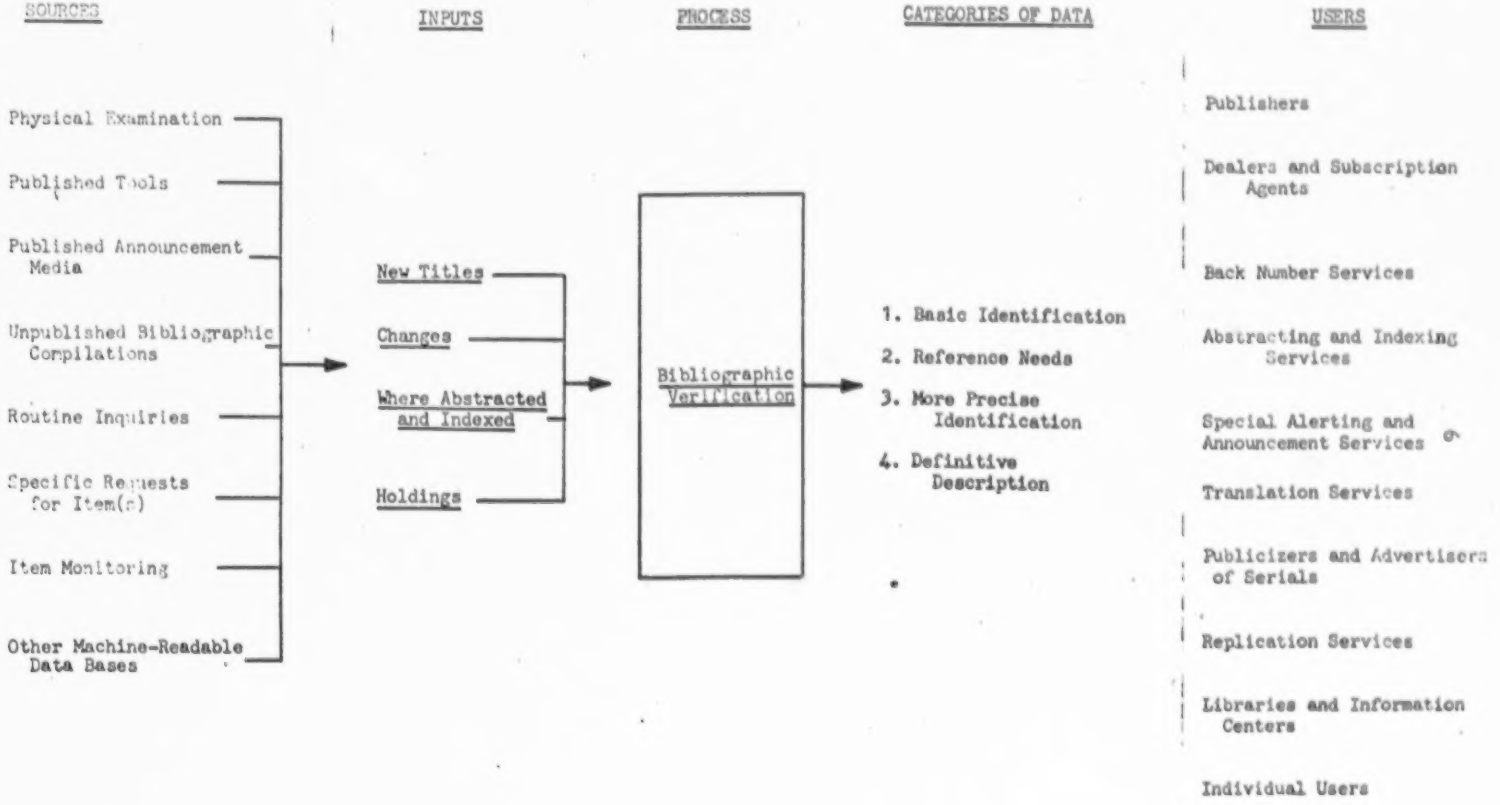

SOURCES
INPUTS
PROCESS
CATEGORIES OF DATA
USERS
Physical Examination
Published Tools
Published Announcement Media
Unpublished Bibliographic Compilations
Routine Inquiries
Specific Requests for Item(s)
Item Monitoring
Other Machine-Readable Data Bases
New Titles
Changes
Where Abstracted and Indexed
Holdings
Bibliographic Verification
1. Basic Identification
2. Reference Needs
3. More Precise Identification
4. Definitive Description
Publishers
Dealers and Subscription Agents
Back Number Services
Abstracting and Indexing Services
Special Alerting and Announcement Services
Translation Services
Publicizers and Advertisers of Serials
Replication Services
Libraries and Information Centers
Individual Users

CHART IC

NEEDED CATEGORIES	USES
Category 1,3	1. Publishers Identifying tools for: - knowing what else is published in his field (subject) - determining need for new serial - registration number - standardization and coordination
Category 1,2,3,4	2. Dealers and Subscription Agents Identifying tools for: - communication bridge between agent and library (faster, more accurate service) - bibliographic description - registration number - standardization and coordination - indication of items not published, variations in title citation, and fairly definitive citation
Category 1	3. Back Number Service - bibliographic identification - holdings and location information (who is in the market for what volumes of what titles) - registration number for standardization and coordination
Category 1,2,3	4. Abstracting and Indexing Services - alert to new titles in their subject area which they may want to index - identify title - accurate and standard bibliographic description - holdings and locating - ability to tell their customer where he can find serial - data for cataloging - check on overlap and duplication in A&I coverage
Category 1	5. Special Alerting and Announcement Services - data on new titles and changes
Category 1,2,3	6. Translation Services - bibliographic description and availability of translations
Category 1	7. Publicizers and Advertisers of Serials (Disseminators of new information about new products and services) data on new titles and changes

Category 1

8. Replication Services (e.g. microfilm edition services, photocopy services, reprint publishers)

   Availability and location of items to be replicated. Indication of replication needs through holdings. Registration Number.

Category 1,2,3,4

9. Libraries and Information Centers
   - alerting service on new titles, where available, description information, acquisition directories (for acquisitions librarians)
   - alerting service on new titles by subject, subject bibliographies (for selection librarians)
   - complete bibliographic description, full cataloging data (for serials catalogers)
   - to use for information about serials, statistical counts on population, etc.

   Administrators and Information System Designers (general library use)
   - Identification and location information for inter-library loan
   - Machine-readable information for input to local serials processing
   - Acquisition tools
   - Cataloging data
   - Special listings
   - Coordination and Standardization

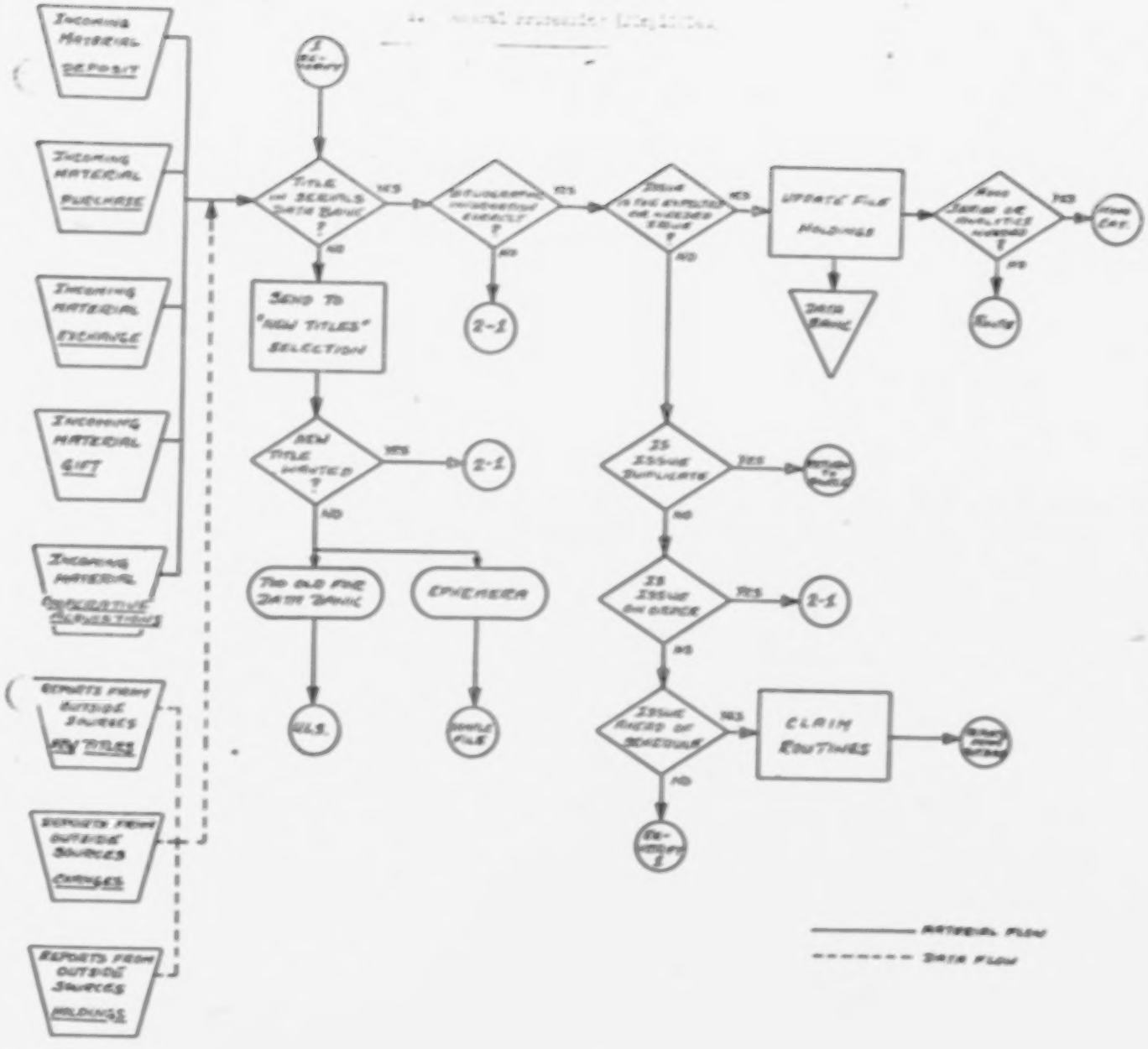
Category 1,2,3

10. Individual Users (e.g. authors, researchers, reference librarians

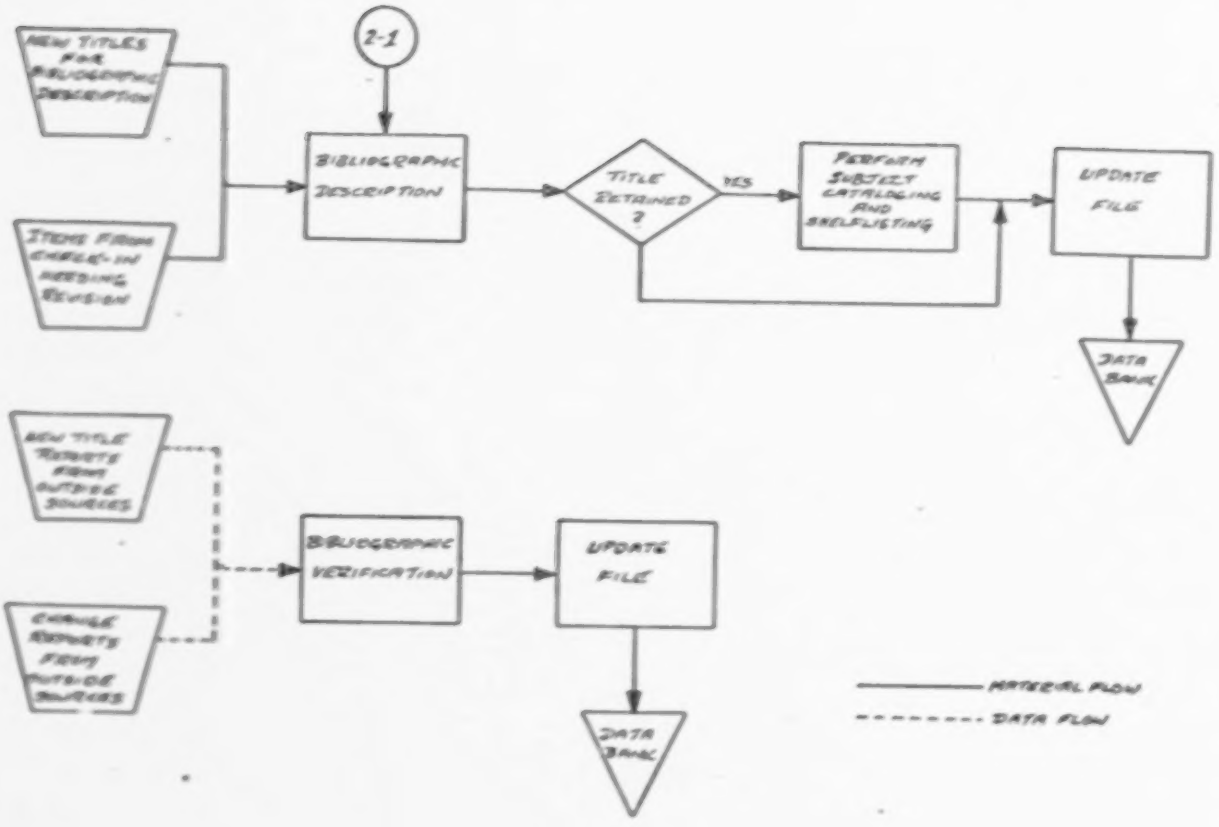
    - Identifying and locating tools
    - Retrieval from title by any selected data element, e.g. language, past and present variations in title
    - Subject bibliographies and guide to subject of individual articles through abstracting and indexing coverage

Chart II shows in simplified form the process necessary to transform the inputs to outputs which will accomplish the stated uses. The process has been simplified in order to highlight the essential logic and to make this logic descriptive of all levels.

CHART II


INCOMING MATERIAL DEPOSIT
INCOMING MATERIAL PURCHASE
INCOMING MATERIAL EXCHANGE
INCOMING MATERIAL GIFT
INCOMING MATERIAL
REPORTS FROM OUTSIDE SOURCES
REPORTS FROM OUTSIDE SOURCES
REPORTS FROM OUTSIDE SOURCES
TITLE OR DETAILS DATA BANK ?
NO
YES
SEND TO "NEW TITLES" SELECTION
NEW TITLE WANTED ?
YES
2-1
NO
THE OLD FILE DATA BANK
EPHEMERA
U.L.S.
SIMPLE FILE
2-1
ISSUE TO THE EXPECTED OR NEEDED ISSUE ?
NO
IS ISSUE DUPLICATE
YES
NO
IS ISSUE ON ORDER
YES
2-1
NO
ISSUE ABEAD OF SCHEDULE
YES
CLAIM ROUTINES
NO
UPDATE FILE HOLDINGS
DATA BANK
MATERIAL FLOW
DATA FLOW

2. Bibliographic Description (Simplified)



SCOPE AND IMPLEMENTATION

The scope of the program should fulfill the maximum number of uses and achieve the most effectiveness with the least amount of cost. Currently published serials represent a portion of the total universe of serials that most nearly fulfills this goal. Whether all serials that are currently published should initially be entered and retained in an on-line basis or whether those serials which have subsequently ceased publication should be moved to secondary storage is a decision that can be made only after detailed system specifications have begun. The latter would achieve a 28% cost per title reduction; however the problems and costs of file linkage must be taken into consideration.

Limiting the scope of the program to a specific subject area would significantly limit the usefulness of the National Serials Data Program since this would limit the potential users to a smaller segment of the information handling community. However, it may be necessary to limit a pilot project by subject in order to obtain a data base of manageable size. Neither would the library and library-related organizations be adequately served if the program were limited to a specific language or languages. The Nelson survey, while indicating some differences of opinion on the definition of a serial, did show that the great majority of potential users defined serial in a broad sense. Therefore, restricting the scope of the program to a specific type of serial, e.g., periodicals, would limit the usefulness of the program. Regarding the inclusion of material on the basis of the serial's substantive content, the decisions presently used in the compilation of New Serial Titles are widely accepted by the information handling community.

THE NATIONAL NETWORK

Once agreement has been reached on a machine format for serial records, a network for exchanging this machine readable information must be created. Primary considerations in the creation of this network are: the assignment of responsibility for capturing, entering, and transmitting bibliographic and control information about serials publications either by type or broad subject category on a national scale; establishment of an efficient method for communicating serials information relating to holdings and accessibility among all members (or nodes) in the network interconnecting the national serials information network; the identification and dedication of transmission facilities; and lastly, the provision within the system of archival responsibility for converting retrospective information about serial publications to compatible machine-readable form and for maintaining non-current machine-readable serials information.

A national serials network can operate most effectively if constructed on hierarchical levels. each level having specific responsibilities and providing specific services and products. Chart III depicts such a system in graphic form.

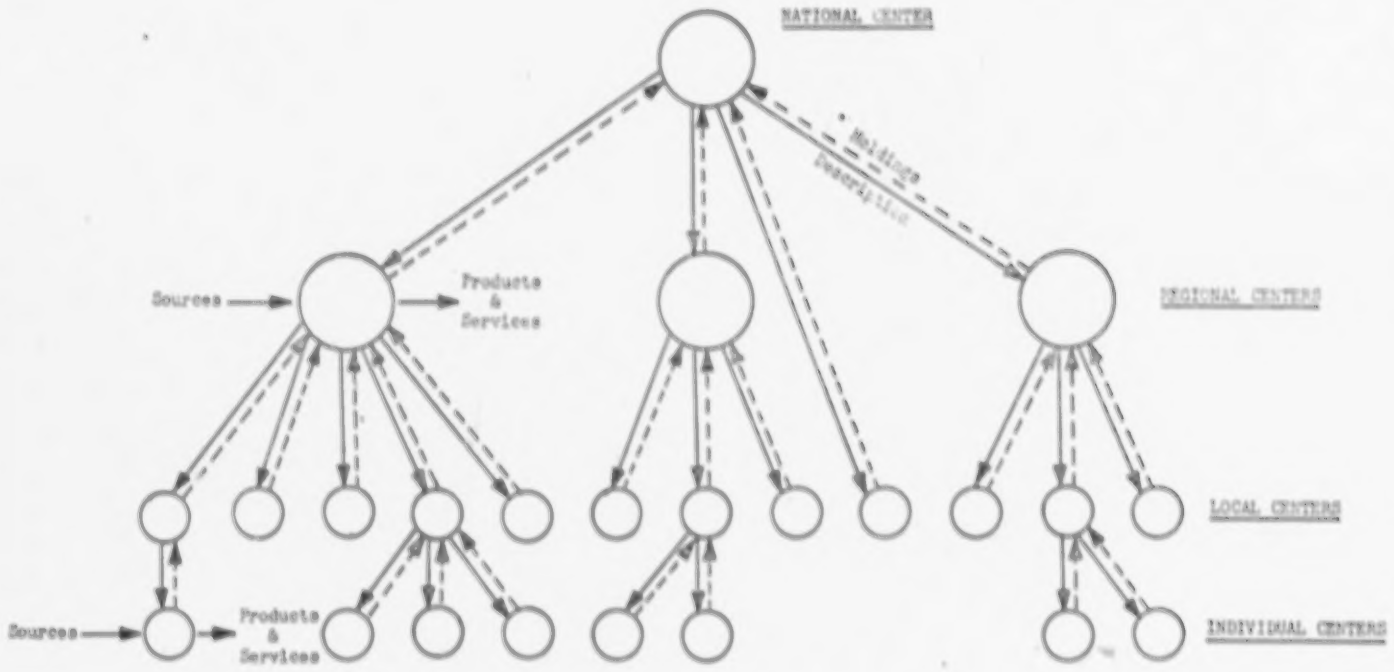
The national center would have one central data bank, composed of categorized files, with linkages to the regional centers and large primary centers. The central data bank will include all entries. Other sites will include entries on a selective basis.

The national center will operate as a collector of information and a wholesaler of information. It will provide predominantly bibliographic information and national control information to regional centers and large primary centers. The flow of input will be bibliographic information flowing down from the national center to regional centers and then to local centers. Some indication of regional holdings will be transmitted from the national center. Similarly, some descriptive information (e.g., items not yet included in the central data bank), will flow from local centers to regional centers and on to the national center. This will serve as an alerting service for the higher echelons. Holdings information will flow from the lowest level upward. Specific holdings will be available only at the lowest level with only summarized holdings being reported to the next higher level.

The products of the national center can be distributed in printed form, on magnetic tape, or via direct access, depending on the ability and willingness of the users to pay the necessary costs. The national center will not be responsible for the costs incurred by the users at any level for memory, communications lines, etc.

CHART II

SUGGESTED NATIONAL NETWORK



Products & Services

- Reports on New Titles
- Data on Changes & Cessations
- National Union List
- Descriptive & Cataloging Information
- Specialized Listings

- Regional Union Lists
- New Titles
- Indication of Holdings

- New Titles
- Detailed Holdings

- New Titles
- Detailed Holdings

NATIONAL CENTER: all titles; indication of region holding the titles
REGIONAL CENTERS: summarized holdings; titles applicable to their region
LOCAL CENTERS: specific holdings; detailed bibliographic description for titles in their area
INDIVIDUAL CENTERS: specific holdings; detailed bibliographic description for titles in their area

Description (with some holdings)

Holdings (with some description)

The population of serials continues to rise even as means are sought for serial control. The difficulty increases exponentially as more titles are added to existing collections. New collections are beginning at the same time that coverage of established ones are expanding to meet interdisciplinary needs. Libraries and library related communities are finding it increasingly difficult to control serial literature. Many have gone to machine assisted approaches in independent efforts. Standards are needed to avoid duplication of effort and hence to reduce costs. It is imperative that the Serials Data Program take the lead in setting these standards. The Program must also move ahead quickly on a pilot project so that experience on living data can be brought to bear on planning for the National Serials Information System.

Appendix I. DATA ELEMENT LIST

INTRODUCTION

Published sources and practicioners of automation in serials control yielded a large number of data elements which had to be considered in a national serials data program. Each of the elements which might be included was represented by a common definition. A definition developed by the program has been taken from an established source, compounded from several definitions and arrived through discussions with those who are currently operating serials systems.

Both "national" and "local" elements have been included. National data elements are those which are intended to be used in many individual systems and which would be required for unambiguous and full communication between centers. The content of each of these would remain constant across centers. Local data elements are those used to control serials within a center. Although the definitions of a local element may be established, content will vary from center to center. Routing information files, for example, apply to all libraries which route serials. The specifics of routing for a given serial or the degree of control will vary from library to library.

The format defined for the national program must include those "national" elements which are necessary to communication and must provide for "local" elements which will be carried with this information in a local system.

TABLE OF CONTENTS

	PAGE
NATIONAL DATA ELEMENTS	
STATUS FIELDS	1
CONTROL NUMBERS	4
CLASSIFICATION NUMBERS	12
BIBLIOGRAPHIC FIELDS	18
IMPRINT	46
INDICATOR FIELDS	53
HOLDINGS AND UNION LIST FIELDS	75
TPI AND INDEX INFORMATION FIELDS	82
FREQUENCY FIELDS	93
LOCAL DATA ELEMENTS	
ACQUISITIONS	97
BINDING	117
ROUTING INFORMATION	132
LOCAL LOCATION FIELDS	136
CLAIMING FIELDS	138